



3 1761 11631976 5

CAI
BS 92
- Z 511



Digitized by the Internet Archive
in 2023 with funding from
University of Toronto

<https://archive.org/details/31761116319765>



Statistics
Canada

Statistique
Canada

92 Z 511

Geography Division

Division de la géographie

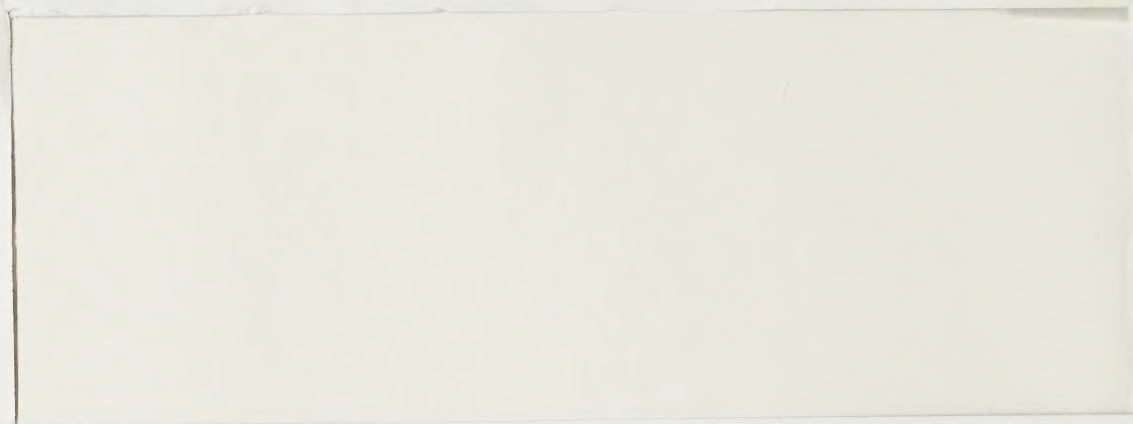
Working Document

Document de travail

Metropolitan Atlas:
1981 Census of Canada
A Procedural Guide

Guide de préparation
des atlas métropolitains
au recensement du Canada de
1981

Canada



Statistics Canada

Statistique Canada

Geography Division

Division de la géographie

Spatial Delineation and
Analysis Section

Section de la délimitation
géographique et de l'analyse

Working Document

Document de travail

Metropolitan Atlas:
1981 Census of Canada
A Procedural Guide

Guide de préparation
des atlas métropolitains
au recensement du Canada de
1981

By

par

Grafton Ross

Grafton Ross

Published under the
authority of the
Minister of Supply and
Services Canada

Publication autorisée par le
ministre des
Approvisionnement et Services
Canada

Minister of Supply
and Services Canada
1984

Ministre des Approvision-
nements et Services Canada
1984

January 1984

Janvier 1984

TABLE OF CONTENTS

Page

1. INTRODUCTION	
1.1 Purpose	5
1.2 General Description	5
1.3 Centres	6
1.4 Themes	6
1.5 Map Production	7
2. DATA DISPLAY	
2.1 General Remarks	8
2.2 Alternative Layouts	8
2.3 Title Box	12
2.4 Dot Distribution Map	12
2.5 Dot Distribution Map Legend	13
2.6 Choropleth Map	13
2.7 Choropleth Map Legend	14
2.8 Frequency Histogram of Relative Values	14
2.9 Comparative Data Table	15
2.10 Cumulative Bar Graph	15
2.11 Frequency Histogram of Absolute Values	16
3. DATA ANALYSIS	
3.1 General Remarks	17
3.2 UNIVARIATE	18
3.3 GDMS	24
3.4 FASTCLUS	27
3.5 DISCRIM	34
3.6 CHANG	41
4. DATA RETRIEVAL	
4.1 General Remarks	48
4.2 Specification of Variables for Retrieval	48
4.3 Specification of Variables for Mapping	49
5. OPERATIONAL PROCEDURES	
5.1 General Remarks	50
5.2 Specification of Variables	52
5.3 Preparation of Text and Comparative Data	54
5.4 Selection of Class Intervals and Dot Rate	58

TABLE DES MATIÈRES

Page

1. INTRODUCTION	
1.1 Objet	5
1.2 Vue d'ensemble	5
1.3 Centres choisis	6
1.4 Thèmes	6
1.5 Production de cartes	7
2. PRÉSENTATION DES DONNÉES	
2.1 Observations générales	8
2.2 Choix du mode de présentation	8
2.3 Titre encadré	12
2.4 Carte de répartition par points	12
2.5 Légende de la carte de répartition par points	13
2.6 Carte choroplèthe	13
2.7 Légende de la carte choroplèthe	14
2.8 Histogramme de fréquence de valeurs relatives	14
2.9 Tableau de données comparatives	15
2.10 Histogramme cumulatif	15
2.11 Histogramme de fréquence de valeurs absolues	16
3. ANALYSE DES DONNÉES	
3.1 Observations générales	17
3.2 UNIVARIATE	18
3.3 GDMS	24
3.4 FASTCLUS	27
3.5 DISCRIM	34
3.6 CHANG	41
4. EXTRACTION DES DONNÉES	
4.1 Observations générales	48
4.2 Détermination des variables aux fins de l'extraction des données	48
4.3 Détermination des variables aux fins de l'établissement des cartes	49
5. MÉTHODES OPÉRATIONNELLES	
5.1 Observations générales	50
5.2 Détermination des variables	52
5.3 Préparation du texte et des données comparatives	54
5.4 Sélection des intervalles de classe et de la densité des points	58

LIST OF FIGURES

	Page
1. Map Layout 1	10
2. Map Layout 2	11
3. Sample UNIVARIATE Output	19
4. Sample GDMS Output	25
5. Sample FASTCLUS Output	28
6. Sample DISCRIM Output	35
7. Sample CHANG Output	42
8. General Work Flow	51
9. Form 1 (Specification of Variables)	53
10. Form 2 (Specification of Text)	55
11. Form 3 (Specification of Comparative Data)	57
12. Form 4 (Specification of Class Intervals and Dot Rate)	59

LISTE DES FIGURES

	Page
1. Mode de présentation n° 1	10
2. Mode de présentation n° 2	11
3. Spécimen de sortie imprimée du programme UNIVARIATE	19
4. Spécimen de sortie imprimée du programme GDMS	25
5. Spécimen de sortie imprimée du programme FASTCLUS	28
6. Spécimen de sortie imprimée du programme DISCRIM	35
7. Spécimen de sortie imprimée du programme CHANG	42
8. Déroulement général des opérations	51
9. Formule 1 (Détermination des variables)	53
10. Formule 2 (Détermination du texte)	55
11. Formule 3 (Détermination des données comparatives)	57
12. Formule 4 (Détermination des intervalles de classe et de la densité des points)	59

1. INTRODUCTION

1.1 Purpose

This guide describes the features of the Metropolitan Atlas Series, the production procedures that have been established for it, and the inputs that will be required from each subject-matter officer in order to produce maps for his/her theme(s).

The main body of the guide is organized in reverse chronological order; that is, it begins with a description of the finished product (Chapter 2 - Data Display) and then moves to the Data Analysis (Chapter 3) and Data Retrieval (Chapter 4) operations. Chapter 5 details the Operational Procedures and describes the various forms that have been developed for submitting subject-matter input.

1.2 General Description

A metropolitan atlas is a catalogued bulletin containing a large number of maps and graphs which illustrate a variety of census themes, or variables, for one census metropolitan area (CMA). Both 2A and 2B data are presented in a single atlas (i.e. data collected from all households and from a 20% sample of households). Mapping is at the census tract (CT) level. The atlas is designed as a complement to the CT Profile bulletins and reveals spatial patterns in the data not immediately apparent when presented in tabular form.

Thematic map atlases are widely used in urban planning, social work, and education - primarily to gain general impressions of the data which can then be followed up with more detailed study. The themes selected for presentation in the Metropolitan Atlas Series are derived from a survey of these users, following the release of a prototype atlas in 1979, and an analysis of similar atlas products in print.

1. INTRODUCTION

1.1 Objet

Le présent guide décrit les caractéristiques de la Série d'atlas métropolitains, les méthodes de production qui ont été mises au point à cette fin ainsi que tout ce que chaque agent spécialisé doit faire pour produire les cartes en fonction du thème choisi.

La partie principale du guide suit un ordre chronologique inversé. En effet, elle débute par une description du produit fini (chapitre 2 - Présentation des données) et traite ensuite la question des opérations d'analyse des données (chapitre 3) et d'extraction des données (chapitre 4). Le chapitre 5 explique en détail les méthodes opérationnelles et décrit les diverses formules qui ont été conçues pour permettre aux agents de soumettre l'information spécialisée à traiter.

1.2 Vue d'ensemble

Un atlas métropolitain est un bulletin inscrit au catalogue dans lequel on trouve un nombre important de cartes et de graphiques servant à illustrer une variété de thèmes du recensement ou de variables pour une région métropolitaine de recensement donnée (RMR). On présente dans le même atlas des données produites à partir des questionnaires 2A et 2B (c'est-à-dire les données recueillies auprès de l'ensemble des ménages et auprès d'un échantillon formé de 20 % des ménages). Les cartes sont établies au niveau des secteurs de recensement (SR). Les atlas visent à compléter les bulletins de la Série des profils des SR et révèlent des caractéristiques spatiales qui ne peuvent être perçues directement lorsque les données sont présentées sous forme de tableaux.

Les atlas de cartes thématiques sont d'une grande utilité dans les domaines de l'urbanisme, du travail social et de l'éducation. Ils permettent essentiellement aux utilisateurs de situer les résultats dans une perspective globale avant de procéder, au besoin, à des études plus détaillées. Pour le choix des thèmes présentés dans la Série d'atlas métropolitains, on s'est fondé sur les résultats d'une enquête menée auprès de ces utilisateurs à la suite de la publication d'un prototype d'atlas en 1979, ainsi que sur une analyse d'atlas semblables déjà imprimés.

1.3 Centres

The series is comprised of 12 atlases - one for each of the following CMAs:

St. John's	Ottawa-Hull	Regina
Halifax	Toronto	Calgary
Québec	Hamilton	Edmonton
Montréal	Winnipeg	Vancouver

These particular centres were selected on the basis of their size and importance within their respective regions. Included in the series are all CMAs over 500,000 population, plus three smaller CMAs. Of the 12 provinces and territories, only New Brunswick, Prince Edward Island, Yukon and Northwest Territories are not represented in the series.

Like the CF Profile bulletins, each atlas is designed to stand alone. While the common selection of themes for all 12 centres will permit inter-CMA comparisons, surveys have indicated that most users are primarily interested in their own CMA. Therefore, when selecting class intervals and other key values for mapping, subject-matter officers should bear in mind that appropriateness to each particular CMA is more important than standardization across the series.

1.4 Themes

The following preliminary themes have been selected for presentation in the Metropolitan Atlas Series:

POPULATION

- (1) Population distribution, 1981
- (2) Population change, 1976-1981

AGE GROUPS

- (3) Under 5 years
- (4) 5-19 years
- (5) 65 years and over

MOTHER TONGUE

- (6) English
- (7) French
- (8) Predominant non-official language(s)

EDUCATION

- (9) Less than Grade 9
- (10) Post-secondary

OCCUPIED PRIVATE DWELLINGS

- (11) Owned Dwellings
- (12) Constructed in 1960 or before
- (13) Requiring major repair
- (14) Average value (owned dwellings)

1.3 Centres choisis

La série comprend 12 atlas, un pour chacune des RMR suivantes:

St. John's	Ottawa-Hull	Regina
Halifax	Toronto	Calgary
Québec	Hamilton	Edmonton
Montréal	Winnipeg	Vancouver

Ces centres ont été choisis à cause de leur taille et de leur importance à l'intérieur de leur région respective. On retrouve dans la série toutes les RMR qui comptent plus de 500,000 habitants et trois autres RMR plus petites. Sur l'ensemble des 12 provinces et territoires, seuls le Nouveau-Brunswick, l'Île-du-Prince-Édouard, le Yukon et les Territoires du Nord-Ouest ne sont pas représentés dans la série.

Comme dans le cas des bulletins de la Série des profils des SR, chaque atlas est conçu de façon indépendante. S'il est vrai que le choix de thèmes communs aux 12 centres permettra d'établir des comparaisons entre les RMR, des enquêtes ont révélé que la plupart des utilisateurs s'intéressent avant tout à leur propre RMR. Par conséquent, au moment de déterminer les intervalles de classe et autres valeurs essentielles à l'établissement des cartes, les agents spécialisés doivent se préoccuper davantage des besoins particuliers à chaque RMR que de l'uniformisation de la série.

1.4 Thèmes

Voici les principaux grands thèmes qui sont inclus dans la Série d'atlas métropolitains:

POPULATION

- (1) Répartition de la population, 1981
- (2) Variation de la population, 1976-1981

GROUPES D'ÂGE

- (3) Moins de 5 ans
- (4) 5-19 ans
- (5) 65 ans et plus

LANGUE MATERNELLE

- (6) Anglais
- (7) Français
- (8) Autre(s) langue(s) prédominante(s)

SCOLARITÉ

- (9) N'ayant pas atteint la 9^e année
- (10) Niveau postsecondaire

LOGEMENTS PRIVÉS OCCUPÉS

- (11) Logements Possédés
- (12) Construction datant de 1960 ou avant
- (13) Exigeant des réparations importantes
- (14) Valeur moyenne (logements occupés par leur propriétaire)

PRIVATE HOUSEHOLDS

- (15) Shelter cost (owned)
- (16) Shelter cost (rented)
- (17) Length of occupancy
- (18) Number of persons per household
- (19) Households with one person
- (20) Households with families
- (21) Households with lone-parent families

LABOUR FORCE ACTIVITY

- (22) Participation rate (males)
- (23) Participation rate (females)
- (24) Unemployment rate
- (25) Managerial, Scientific and related occupations

PLACE OF WORK

- (26) Working in the central business district
- (27) Working in the census tract

INCOME

- (28) Total income (private households)
- (29) Total income (census families)
- (30) Total income (non-family persons)
- (31) Low income families and unattached individuals

Each of these general themes will be developed in detail by the individual officer who must specify precisely what is to be shown on each map and graph for his/her theme, and must also satisfy himself/herself that the data appear reasonable.

All themes for the Metropolitan Atlas Series are based on data retrieved from the 2A and 2B Profile user summary tapes at the CT level.

1.5 Map Production

The Metropolitan Atlas Series is a high technology, state-of-the-art product. Significant advances in data analysis procedures, in computer mapping software, and in plotting hardware have been brought together for the first time to produce these atlases. The Geocartographics Sub-division (Systems and Data Processing Branch) and the Geography Division have worked many months to develop and test this production system.

MÉNAGES PRIVÉS

- (15) Prix du logement (possédé)
- (16) Prix du logement (loué)
- (17) Durée d'occupation
- (18) Nombre de personnes par ménage
- (19) Ménages constitués d'une personne
- (20) Ménages composés de familles
- (21) Ménages composés de familles monoparentales

ACTIVITÉ

- (22) Taux d'activité (hommes)
- (23) Taux d'activité (femmes)
- (24) Taux de chômage
- (25) Administrateurs, Directeurs, Scientifiques et personnels connexes

LIEU DE TRAVAIL

- (26) Travaillant dans le secteur du centre des affaires
- (27) Travaillant dans le secteur de recensement

REVENU

- (28) Revenu total (ménages privés)
- (29) Revenu total (familles de recensement)
- (30) Revenu total (personnes hors famille)
- (31) Familles et personnes seules à faible revenu

Chacun de ces grands thèmes est élaboré en détail par l'agent responsable qui doit préciser le contenu de chaque carte et graphique correspondant au thème choisi, et aussi s'assurer de la pertinence des données.

Tous les thèmes de la Série d'atlas métropolitains sont basés sur les données extraites des profils 2A et 2B sous forme de bandes sommaires au niveau des SR.

1.5 Production de cartes

La Série d'atlas métropolitains est un produit d'avant-garde qui fait appel à une haute technologie. Pour produire ces atlas, on a appliqué pour la première fois l'ensemble des progrès les plus importants réalisés dans le domaine des méthodes d'analyse des données, des programmes informatiques de cartographie et du matériel de traçage. La Sous-division de la géocartographie (Direction des systèmes informatiques et du traitement des données) et la Division de la géographie ont consacré de nombreux mois à l'élaboration et à la mise au point de ce système de production.

Data analysis procedures are described in detail in Chapter 3 of this guide. Essentially, what has been accomplished here is the linking together of a series of general-purpose analytical programs with others developed specifically for mapping applications, into a single operational package.

Software support for computer mapping is provided by GIMMS (Geographic Information Manipulation and Mapping System), developed by Professor T. Waugh at the University of Edinburgh. Statistics Canada uses Version 4 of this highly sophisticated mapping system.

The atlases are produced entirely on a GERBER 4442 high-speed drum plotter. This device uses a special photohead assembly in which a focused beam of light passes over a sheet of photographic film mounted on the drum. Pages are plotted at actual size so that the film can be sent directly for printing of the bulletins with no intermediate steps or loss of quality.

Map shading is accomplished with a series of dot screen and line cross-hatching patterns developed by the Geocartographics Sub-division. Page layouts were designed by Geocartographics under the direction of the Geography Division. Census tract boundaries were encoded by the Census Geographic Master File and enhanced to cartographic standards by Geocartographics.

2. DATA DISPLAY

2.1 General Remarks

The atlases combine maps, graphs and text to present each theme in as thorough and unbiased a manner as possible. The individual elements of this data display, illustrated on the sample layouts (Figures 1 and 2) and described below, have been developed through user surveys of our prototype bulletin and analysis of other similar atlas products in print. The Metropolitan Atlas Series will be far more comprehensive than any similar product reviewed to date.

2.2 Alternative Layouts

The main display elements are the maps. There are two for each theme - a dot distribution map to show absolute values

Les méthodes d'analyse des données sont décrites en détail au chapitre 3. Essentiellement, on a réussi à créer un ensemble opérationnel unique en reliant plusieurs programmes généraux d'analyse à d'autres spécialement destinés à des travaux de cartographie.

La cartographie informatique s'appuie sur l'aide logicielle fournie par le GIMMS (Système de traitement des données géographiques et cartographiques) conçu par T. Waugh, professeur à l'"University of Edinburgh". Statistique Canada utilise la version 4 de ce système de cartographie très perfectionné.

Les atlas sont produits entièrement sur un GERBER 4442, traceur à tambour rapide. Cette unité utilise une tête photographique spéciale dans laquelle un trait lumineux pouvant être grossi ou diminué à volonté passe au-dessus d'un film sensible monté sur le tambour. Les cartes sont tracées selon le format réel; le film peut donc être envoyé directement au service d'impression des bulletins sans autres étapes intermédiaires ou perte de qualité.

Le tramage des cartes se fait à l'aide d'une série de pointillés et de lignes hachurées, selon des modèles mis au point spécialement par la Sous-division de la géocartographie. La mise en page a été conçue par le personnel de la Sous-division de la géocartographie avec l'aide de celui de la Division de la géographie. Les limites des secteurs de recensement, enregistrées en code dans le fichier géographique principal du recensement, ont été modifiées par le personnel de la géocartographie en fonction des normes cartographiques en vigueur.

2. PRÉSENTATION DES DONNÉES

2.1 Observations générales

Les atlas contiennent des cartes, des graphiques et un texte explicatif de sorte que chaque thème est présenté de la manière la plus complète et la plus fidèle possible. Les différents éléments qui entrent dans cette présentation des données sont illustrés dans les modèles (figures 1 et 2) décrits ci-dessous et sont fondés sur les résultats de l'enquête menée auprès des utilisateurs à l'aide de notre bulletin prototype, ainsi que sur l'analyse d'autres atlas semblables déjà produits. La Série d'atlas métropolitains sera beaucoup plus compréhensive que tout autre produit du même genre analysé à ce jour.

2.2 Choix du mode de présentation

Les cartes constituent les principaux éléments de la présentation. Il en existe deux types pour chaque thème présenté: une carte de répartition

(e.g., number of persons under 5 years of age) and a choropleth map to show relative values (e.g., percent of persons under 5 years). For most themes, such as age groups, a direct correspondence exists between the data displayed on the dot distribution and choropleth maps; for some themes (e.g., dwelling value, shelter cost, household size, income) the two maps are related but not necessarily identical. Two layouts have been developed to accommodate these differences between what may be called "qualitative" and "quantitative" variables.

Layout 1 (for the "qualitative" themes) uses a cumulative bar graph to set the variable in context (e.g., the under-5 group in relation to other age groups) and to compare the CMA with the province and the country as a whole. For the "quantitative" themes, where this bar graph cannot be applied, Layout 2 presents in its place a frequency histogram showing the distribution of CIs by absolute values. In all other respects the two layouts are identical.

par points pour les valeurs absolues (par exemple, le nombre de personnes âgées de moins de 5 ans) et une carte choroplèthe pour les valeurs relatives (par exemple, la proportion de personnes âgées de moins de 5 ans). Pour la plupart des thèmes, notamment celui des groupes d'âge, on peut établir une correspondance directe entre les données présentées sur la carte de répartition par points et celles de la carte choroplèthe. Pour certains thèmes (par exemple, la valeur du logement, le coût du logement, la taille des ménages et le revenu), les deux types de cartes sont connexes mais non nécessairement identiques. Deux modes de présentation ont été conçus pour tenir compte des différences entre ce qu'on pourrait appeler les variables "qualitatives" et les variables "quantitatives".

Pour le mode de présentation n° 1 (dans le cas des thèmes "qualitatifs"), on utilise un histogramme cumulatif qui permet de situer la variable dans son contexte (par exemple, le groupe des personnes âgées de moins de 5 ans par rapport aux autres groupes d'âge) et de comparer les chiffres de la RMR avec ceux de la province et de l'ensemble du pays. Lorsqu'on ne peut avoir recours à ce type d'histogramme pour l'analyse de thèmes "quantitatifs", on utilise à la place un histogramme de fréquence (mode de présentation n° 2) qui montre la distribution des SR en chiffres absolus. À tous les autres points de vue, les deux modes de présentation sont identiques.

99-125

POPULATION UNDER FIVE YEARS

The number of persons under five years of age is shown. This age group corresponds roughly to the pre-school population.

POPULATION DE MOINS DE CINQ ANS

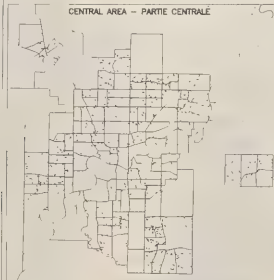
On présente le nombre de personnes moins de cinq ans. Ce groupe d'âge correspond approximativement à la population d'âge préscolaire.

EDMONTON

CENSUS TRACTS - SECTEURS DE RECENSEMENT

CENTRAL AREA - PARTIE CENTRALE

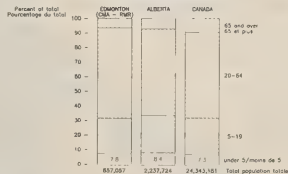
CMA - RMR



One dot represents 50 persons
Un point représente 50 personnes

□ population <25

POPULATION BY AGE GROUP POPULATION SELON LE GROUPE D'ÂGE



SOURCE: 1981 CENSUS OF CANADA
PRODUCED BY STATISTICS CANADA

SOURCE: RECENSEMENT DU CANADA 1981
FAIT PAR STATISTIQUE CANADA

99-925

POPULATION UNDER FIVE YEARS

Persons under five years of age are expressed as a percentage of the total population.

POPULATION DE MOINS DE CINQ ANS

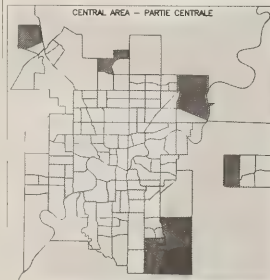
Le nombre de personnes moins de cinq ans est exprimé en pourcentage de la population totale.

EDMONTON

CENSUS TRACTS - SECTEURS DE RECENSEMENT

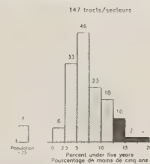
CENTRAL AREA - PARTIE CENTRALE

CMA - RMR



17.5% or more/ou plus
12.5% to/à 17.5%
7.5% to/à 12.5%
2.5% to/à 7.5%
less than/moins de 2.5%
□ population <25

NUMBER OF TRACTS BY PERCENT UNDER FIVE YEARS NOMBRE DE SECTEURS SELON LE POURCENTAGE DE MOINS DE CINQ ANS



COMPARATIVE FIGURES CHIFFRES COMPARATIFS

EDMONTON (CMA - RMR) 7.8%
ALBERTA 8.4%
CANADA 7.3%

SOURCE: 1981 CENSUS OF CANADA
PRODUCED BY STATISTICS CANADA

SOURCE: RECENSEMENT DU CANADA DE 1981
FAIT PAR STATISTIQUE CANADA

POPULATION DISTRIBUTION, 1981

The total 1981 population in each census tract is shown. Total populations are shown by dots representing absolute numbers of persons.

RÉPARTITION DE LA POPULATION, 1981

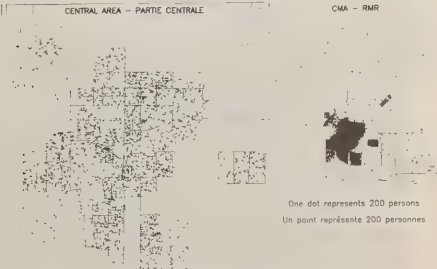
On présente la population totale en 1981 de chaque secteur de recensement. La population totale est illustrée par des points qui représentent le nombre absolu de personnes.

EDMONTON

CENSUS TRACTS - SECTEURS DE RECENSEMENT

CENTRAL AREA - PARTIE CENTRALE

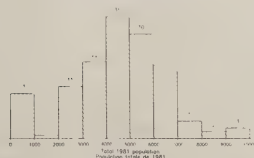
CMA - RMR



One dot represents 200 persons
Un point représente 200 personnes

NUMBER OF TRACTS BY TOTAL POPULATION NOMBRE DE SECTEURS SELON LEUR POPULATION TOTALE

147 Tracts



SOURCE: 1981 CENSUS OF CANADA
PRODUCED BY STATISTICS CANADA

SOURCE: RECENSEMENT DU CANADA DE 1981
ÉLABORÉ PAR STATISTIQUE CANADA

POPULATION CHANGE, 1976 - 1981

Population change between 1976 and 1981 is expressed as a percentage of the total 1976 population for each census tract. Census tracts with no population in 1976 or 1981 are excluded.

VARIATION DE LA POPULATION, 1976 - 1981

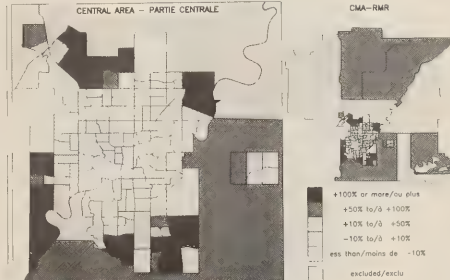
La variation de la population entre 1976 et 1981 est exprimée en pourcentage de la population totale de 1976 pour chaque secteur de recensement. Les secteurs de recensement sans population en 1976 ou en 1981 sont exclus.

EDMONTON

CENSUS TRACTS - SECTEURS DE RECENSEMENT

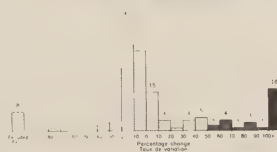
CENTRAL AREA - PARTIE CENTRALE

CMA - RMR



NUMBER OF TRACTS BY PERCENTAGE CHANGE NOMBRE DE SECTEURS SELON LE TAUX DE VARIATION

147 Tracts



SOURCE: 1981 CENSUS OF CANADA
PRODUCED BY STATISTICS CANADA

COMPARATIVE FIGURES CHIFFRES COMPARATIFS

EDMONTON	18.1%
(CMA - RMR)	
ALBERTA	21.7%
CANADA	9.9%

SOURCE: RECENSEMENT DU CANADA DE 1981
ÉLABORÉ PAR STATISTIQUE CANADA

2.3 Title Box

POPULATION UNDER FIVE YEARS

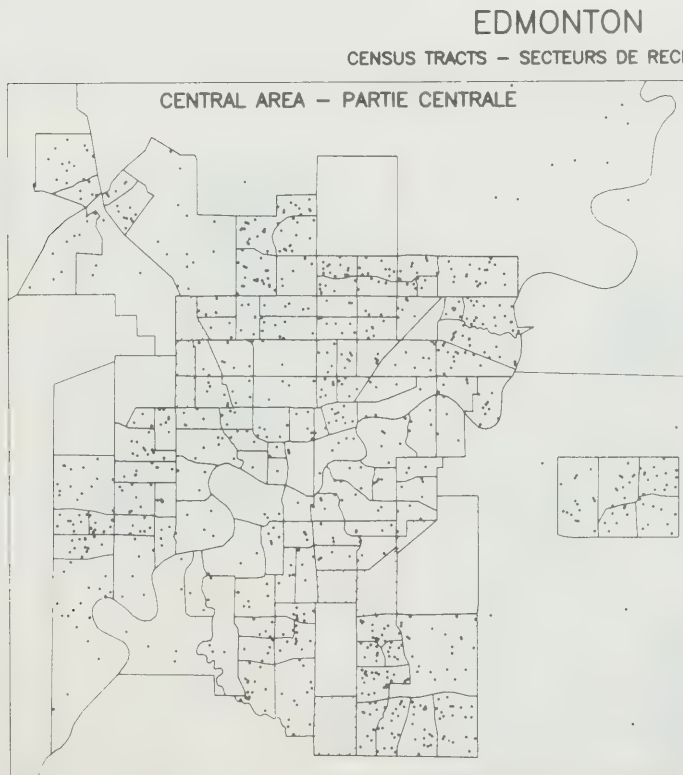
The number of persons under five years of age is shown. This age group corresponds roughly to the pre-school population.

The title box at the top of each page contains two elements - a brief title and a short explanatory text.

The title is limited to two lines of 27 characters each (including spaces) in English and two lines of 30 characters each (including spaces) in French. As such, it serves only to identify rather than explain the theme (e.g., "ENGLISH MOTHER TONGUE", rather than "PERCENT OF THE POPULATION WITH ENGLISH MOTHER TONGUE" or "NUMBER OF PERSONS WITH ENGLISH MOTHER TONGUE").

The detailed explanation of the theme is provided in the text immediately below the title. The explanatory text is limited to six lines of 50 characters each (including spaces) in English and six lines of 55 characters each (including spaces) in French. The text specifies precisely how the particular rate or percentage portrayed on the map was calculated and indicates what criteria for data suppression have been applied, if other than the standard "total population less than 25" criterion. Other brief notes concerning the theme may be included at the subject-matter officer's discretion.

2.4 Dot Distribution Map



2.3 Titre encadré

POPULATION DE MOINS DE CINQ ANS

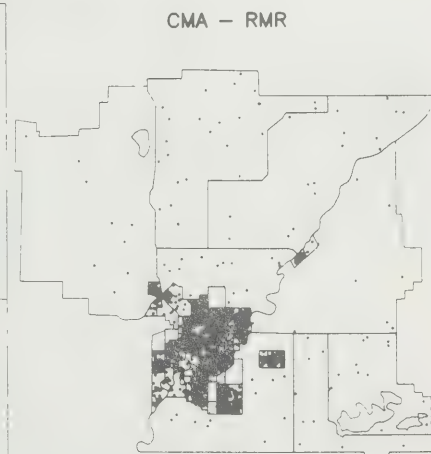
On présente le nombre de personnes moins de cinq ans. Ce groupe d'âge correspond approximativement à la population d'âge préscolaire.

Le titre encadré, au haut de chaque page, comprend un bref titre et un court texte explicatif.

Le titre anglais ne doit pas dépasser deux lignes de 27 caractères chacune (espaces compris) et le titre français, deux lignes de 30 caractères chacune (espaces compris). Il sert ainsi à énoncer plutôt qu'à expliquer le thème (par ex., "ANGLAIS-LANGUE MATERNELLE", plutôt que "POURCENTAGE DE LA POPULATION DONT L'ANGLAIS EST LA LANGUE MATERNELLE" ou "NOMBRE DE PERSONNES DONT L'ANGLAIS EST LA LANGUE MATERNELLE").

L'explication détaillée du thème est fournie dans le texte qui suit immédiatement le titre. Le texte explicatif anglais comprend tout au plus six lignes de 50 caractères chacune (espaces compris) et le texte explicatif français, six lignes de 55 caractères chacune (espaces compris). Le texte précise de quelle façon on a calculé le taux ou le pourcentage qui figure sur la carte et mentionne les normes appliquées à l'élimination de données, si elles diffèrent de la norme courante de "population totale inférieure à 25". L'agent spécialisé peut inclure quelques brefs commentaires sur le thème s'il le juge nécessaire.


2.4 Carte de répartition par points



The dot distribution map portrays the theme in terms of absolute values - each dot representing a fixed number of occurrences (persons, dwellings, etc.). Depending on the number and sizes of CTs in the CMA, the map will be presented either as a single map of the entire CMA or as a CMA map plus one or more insets.

2.5 Dot Distribution Map Legend

One dot represents 50 persons
Un point représente 50 personnes

 population <25

The legend indicates the dot rate used on the map. This rate is established by the subject-matter officer, assisted by a computer program which calculates an "optimum" rate based on the number of occurrences per CT, the land area of each CT, and the physical size of the dots.

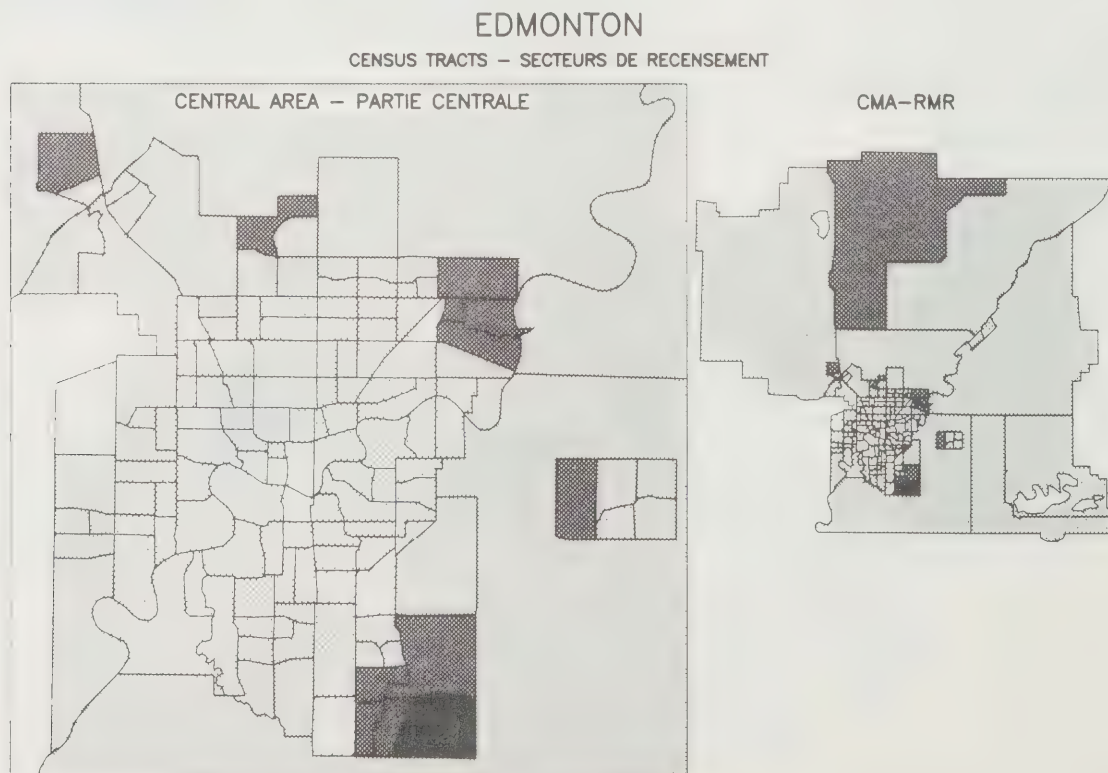
2.6 Choropleth Map

La carte de répartition par points illustre le thème en valeurs absolues, chaque point représentant un nombre fixe d'observations (personnes, logements, etc.). Dépendant du nombre et de la taille des SR situés dans la RMR, il s'agit d'une seule carte de la RMR toute entière ou d'une carte de la RMR complétée par un ou plusieurs cartons.

2.5 Légende de la carte de répartition par points

La légende précise la densité des points qui figurent sur la carte. Pour déterminer cette densité, l'agent spécialisé se fonde sur un programme informatique qui calcule la densité "optimale" en fonction du nombre d'observations par SR, de la superficie de chaque SR et de la grosseur des points.

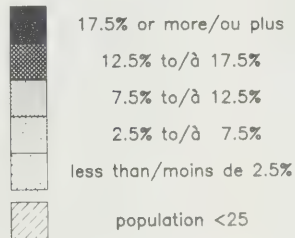
2.6 Carte choroplèthe



The choropleth map portrays the theme in terms of relative values (percentages, rates, etc.) - each shade representing a particular range of values.

The choropleth map and insets are presented at the same scales as the dot distribution map and insets.

2.7 Choropleth Map Legend



The legend indicates the range of values associated with each shade on the map. The subject-matter officer has the option of selecting either three, four or five classes to cover the total range of values for each theme in each centre. In this, and in the selection of the class "breaks" or division points between classes, the officer is assisted by a series of computer programs which cluster CIs into groups based on an analysis of the data.

Note that the class intervals selected for the choropleth map must be multiples of the class intervals selected for the frequency histogram (Item 2.8).

2.8 Frequency Histogram of Relative Values

La carte choroplèthe illustre le thème en valeurs relatives (pourcentages, taux, etc.), chaque hachure représentant une étendue de valeurs particulière.

Pour la carte choroplèthe et ses cartons, on utilise les mêmes échelles que pour la carte de répartition par points et ses cartons.

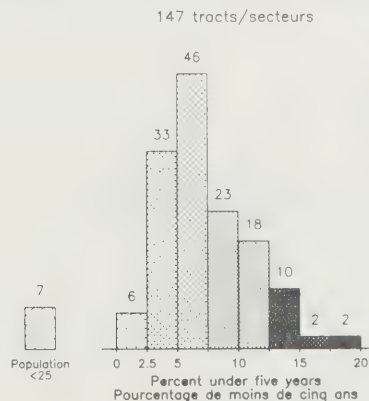
2.7 Légende de la carte choroplèthe

La légende indique l'amplitude des valeurs correspondant à chaque hachure sur la carte. L'agent spécialisé peut choisir trois, quatre ou cinq classes pour représenter l'étendue totale des valeurs de chaque thème dans chaque centre. À cette fin, ainsi que pour la sélection des intervalles ou séparations de classe, l'agent se sert d'une série de programmes informatiques qui répartissent les SR en grappes après analyse des données.

À noter que les intervalles de classe choisis pour la confection de cartes choroplèthes doivent être des multiples des intervalles de classe utilisés dans l'histogramme de fréquence (voir point 2.8).

2.8 Histogramme de fréquence de valeurs relatives

NUMBER OF TRACTS BY PERCENT UNDER FIVE YEARS
NOMBRE DE SECTEURS SELON LE POURCENTAGE DE MOINS DE CINQ ANS



The histogram shows the frequency distribution of CTs across the total range of relative values. The histogram also shows the number of CTs for which data have been suppressed. These same CTs are identified by a special shading pattern on the maps.

The subject-matter officer selects the values for the x-axis of the histogram, assisted by the computer output provided for Item 2.7, so as to yield between 10 and 20 equal-interval classes (e.g., 0% to 100% by 5% intervals). These classes can be regrouped to form the larger classes on the choropleth map, as indicated by the common shading patterns on both.

The histogram title is limited to one line of 62 characters (including spaces) in each of English and French.

2.9 Comparative Data Table

COMPARATIVE FIGURES CHIFFRES COMPARATIFS

EDMONTON (CMA - RMR)	7.8%
ALBERTA	8.4%
CANADA	7.3%

This small table provides comparative CMA, province and Canada figures (supplied by the subject-matter officer) to accompany the choropleth map.

2.10 Cumulative Bar Graph (LAYOUT 1 ONLY)

Cet histogramme indique la distribution de fréquence des SR dans l'amplitude totale des valeurs relatives. Il précise également le nombre des SR pour lesquels des données ont dû être éliminées. Ces SR sont désignés sur les cartes à l'aide d'un motif hachuré distinct.

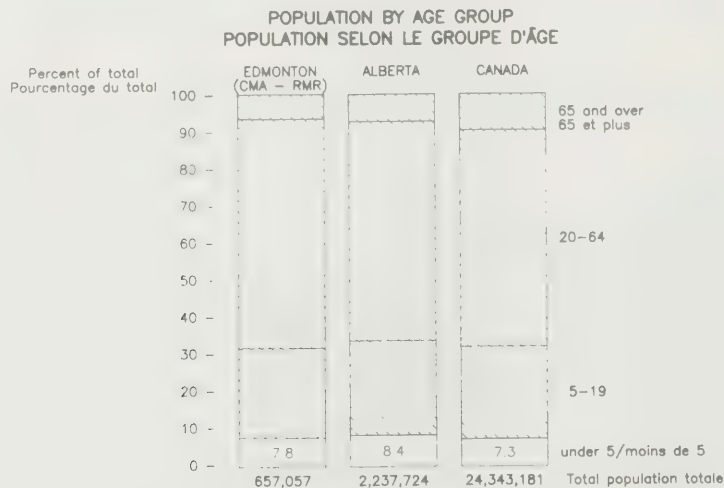
L'agent spécialisé choisit les valeurs devant figurer sur l'axe des x de l'histogramme à l'aide des résultats du traitement par ordinateur mentionné au point 2.7, de façon à obtenir de 10 à 20 classes d'intervalles égaux (par ex., de 0 % à 100 % par intervalle de 5 %). Ces classes peuvent être regroupées pour correspondre aux classes plus grandes de la carte choroplèthe, à l'aide de motifs hachurés communs.

Le titre de l'histogramme ne doit pas excéder une ligne de 62 caractères (espaces compris), tant dans la version anglaise que française.

2.9 Tableau de données comparatives

Ce petit tableau joint à la carte choroplèthe contient des chiffres comparatifs de la RMR, de la province et du Canada, qui sont fournis par l'agent spécialisé.

2.10 Histogramme cumulatif (MODE DE PRÉSENTATION N° 1 SEULEMENT)



The subject-matter officer will determine whether the cumulative bar graph can be applied to his/her theme. If not, the officer will use Item 2.11 instead.

The cumulative bar graph serves two functions: to place the chosen theme into its internal context (e.g., the under-5 group in relation to other age groups) and to compare the CMA with the province and the country as a whole. Three bars of equal height represent the universe from which the theme was selected (e.g., total population) for the CMA, the province and Canada. The bars can be subdivided into from two to five major variable classifications, as selected by the subject-matter officer, and are presented in some logical order (e.g., as used in the CI Profile bulletins). Where these are not collectively exhaustive, the bars are capped with an "other" category. The chosen theme is highlighted and comparative figures for it are placed in the three bars. Comparative universe totals are placed below the three bars.

Where two or more themes have been selected from the same universe for inclusion in the atlas (e.g., the three age groups) the cumulative bar graphs for each will be identical except for the highlighting.

Subject-matter input for this item consists of:

- (1) the title, limited to one line of 62 characters (including spaces) in each of English and French;
- (2) the major variable classifications in English and French, and in the desired order; and
- (3) rounded counts, by major classification and total, for the CMA, the province and Canada.

2.11 Frequency Histogram of Absolute Values (LAYOUT 2 ONLY)

Il appartient à l'agent spécialisé de déterminer s'il peut ou non utiliser l'histogramme cumulatif pour illustrer son thème. Dans la négative, l'agent doit se reporter au point 2.11.

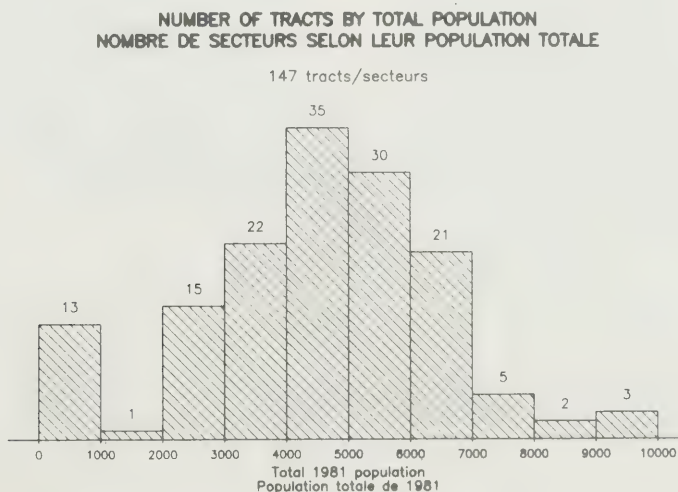
L'histogramme cumulatif remplit deux fonctions: situer le thème choisi dans son contexte interne (par exemple, le groupe des personnes âgées de moins de 5 ans par rapport aux autres groupes d'âge) et comparer les chiffres de la RMR avec ceux de la province et de l'ensemble du pays. Trois barres d'égale hauteur représentent l'univers à partir duquel le thème a été choisi (par exemple, la population totale), une pour la RMR, une pour la province et une autre pour le Canada. Les barres peuvent ensuite être réparties en deux à cinq variables principales choisies par l'agent spécialisé et doivent être présentées dans un certain ordre logique (par exemple, comme c'est le cas dans les bulletins de la Série des profils des SR). Lorsque l'ensemble de ces variables ne forme pas un tout, les barres portent l'en-tête "autres". Le thème sélectionné est mis en relief et les chiffres fournis pour les besoins de comparaison sont inscrits dans les trois barres. Les totaux comparatifs qui représentent l'univers figurent sous les trois barres.

Si l'on décide d'inclure dans l'atlas deux ou plusieurs thèmes à partir du même univers (par exemple les trois groupes d'âge), les histogrammes cumulatifs établis pour chacun d'eux doivent être identiques, sauf pour la partie mise en relief.

Les données d'entrée nécessaires à ce type d'histogramme comprennent:

- (1) le titre, n'excédant pas une ligne de 62 caractères (espaces compris) tant en anglais qu'en français;
- (2) les principales variables, présentées dans les deux langues et dans l'ordre voulu; et
- (3) les chiffres arrondis pour chaque principale variable ainsi que pour les totaux, au niveau de la RMR, de la province et du Canada.

2.11 Histogramme de fréquence de valeurs absolues (MODE DE PRÉSENTATION N° 2 SEULEMENT)



For those themes where Item 2.10 cannot be used, the cumulative bar graph is replaced by a frequency histogram. This histogram is similar to Item 2.8, but shows absolute rather than relative values on the x-axis. The subject-matter officer, assisted by the computer output provided for Item 2.5, selects values so as to yield between 10 and 20 equal-interval classes (e.g., 0 to 5,000, by intervals of 250). Note that the class intervals selected for the histogram must be multiples of the dot rate selected for the dot distribution map (Item 2.5).

The histogram title is limited to one line of 62 characters (including spaces) in each of English and French.

3. DATA ANALYSIS

3.1 General Remarks

A series of computer programs have been developed to assist the subject-matter officer in the selection of appropriate class intervals for the choropleth map and frequency histogram and an appropriate dot rate for the dot distribution map. These programs analyse the CT-level summary tape data to produce "optimum" values for each theme within each individual CMA. The subject-matter officer examines the output of these programs and specifies the actual values which will then be used for displaying his/her theme.

The presentation of data in graphic form is a very powerful means of communication and the impression created is strongly influenced by the particular class "breaks" and other key values selected. For this reason, the subject-matter officer has the ultimate responsibility in their selection, and will modify the optimum values generated by the analytical programs based on his/her knowledge of the theme and on the message he/she wishes to put across.

Dans le cas des thèmes pour lesquels l'histogramme cumulatif ne convient pas, on doit utiliser à la place un histogramme de fréquence de valeurs absolues. Ce type d'histogramme est pratiquement identique à celui décrit au point 2.8, sauf que les valeurs qui figurent sur l'axe des x sont exprimées en chiffres absolus plutôt qu'en chiffres relatifs. L'agent spécialisé se fonde sur les résultats du traitement par ordinateur (voir le point 2.5) pour choisir les valeurs qui lui permettront d'obtenir de 10 à 20 classes d'intervalles égaux (par exemple, de 0 à 5,000, par intervalle de 250). Il importe de rappeler que les intervalles de classe choisis pour la confection de l'histogramme doivent être des multiples de la densité des points déterminée pour la carte de répartition par points (voir le point 2.5).

Le titre de cet histogramme ne doit pas excéder une ligne de 62 caractères (espaces compris) tant dans la version anglaise que française.

3. ANALYSE DES DONNÉES

3.1 Observations générales

On a mis au point une série de programmes informatiques pour aider l'agent spécialisé à déterminer les intervalles de classe appropriés pour la carte choroplèthe et l'histogramme de fréquence ainsi que la densité des points convenant à la carte de répartition par points. Ces programmes analysent les données des bandes sommaires réparties au niveau des SR en vue de déterminer les valeurs "optimales" de chaque thème dans les différentes RMR. L'agent spécialisé examine l'information de sortie de ces programmes et procède à la sélection des valeurs réelles qu'il utilisera pour présenter son thème.

La présentation des données sous forme de graphique est un outil de diffusion de l'information très efficace dont l'effet produit dépend largement du choix des intervalles de classe et des valeurs-clés. La responsabilité d'un tel choix incombe en dernier lieu à l'agent spécialisé qui peut modifier les valeurs optimales des résultats des programmes d'analyse en fonction de sa connaissance du thème choisi et du message qu'il veut transmettre.

At a minimum, these modifications will involve rounding off the optimum values to whole numbers (e.g., an optimum class interval of "21.4% to 28.9%" might be rounded to "20% to 30%"). In some cases certain particular values might have special meaning with respect to the theme, or the subject-matter officer may wish to facilitate comparisons between related themes in the atlas by establishing some common key values on the different maps.

The above comments notwithstanding, the subject-matter officer should endeavour not to stray too far from the optimum values provided, because they are at least unbiased and do in fact reflect the underlying distribution of the data.

3.2 UNIVARIATE

UNIVARIATE is a SAS procedure (Statistical Analysis System, developed by the SAS Institute in Cary, North Carolina) which provides a wide range of summary statistics on variable distributions (see Figure 3). UNIVARIATE will be applied to all the variables used for each theme. In the "age group under 5" example attached, this includes the input variable "total population"(POP 1981), the input variable "population under 5" (POP 0-4) which is also used for dot mapping, and the derived variable "percentage of persons under 5" (PCPOP 0-4) which is used for choropleth mapping.

These statistics are provided as background information to get a better feel for the data. In analysing this output, particularly for the derived variable, subject-matter officers should bear in mind that the units of observation are CTs, not individual persons, dwellings, or households.

Ces modifications peuvent consister simplement à arrondir à des nombres entiers les valeurs optimales obtenues (par ex., un intervalle de classe optimal de "21.4 % à 28.9 %" est arrondi de "20 % à 30 %"). Il peut arriver que certaines valeurs aient un intérêt particulier relativement au thème ou que l'agent spécialisé désire faciliter les comparaisons entre thèmes connexes présentés dans l'atlas en reproduisant des valeurs-clés communes sur les différentes cartes.

L'agent spécialisé doit cependant éviter de trop s'éloigner des valeurs optimales fournies parce qu'elles ont au moins l'avantage d'être sans biais et d'offrir une juste mesure de la distribution sous-jacente des données.

3.2 UNIVARIATE

L'UNIVARIATE est un programme du SAS (Système d'analyse statistique conçu par le SAS Institute à Cary, en Caroline du Nord) qui fournit une vaste gamme de statistiques sommaires sur la distribution des variables (voir figure 3). Le programme UNIVARIATE peut être appliqué à toutes les variables utilisées pour chacun des thèmes. Dans l'exemple ci-joint qui porte sur le "groupe d'âge de moins de 5 ans", il s'agit de la variable d'entrée "population totale" (POP 1981), de la variable d'entrée "population âgée de moins de 5 ans" (POP 0-4) qui est également utilisée pour la carte de répartition par points et de la variable dérivée "pourcentage des personnes âgées de moins de 5 ans" (PCPOP 0-4) qui est également utilisée pour la carte choroplèthe.

Ces statistiques sont fournies à titre de renseignements de base en vue d'une meilleure exploitation des données. Lorsque les agents spécialisés analysent ces données de sortie, particulièrement celles qui touchent la variable dérivée, ils ne doivent pas oublier que les unités d'observation sont des SR, et non des personnes, des logements ou des ménages.

UNIVARIATE

VARIABLE=POPI981

MOMENTS

N 147 SUM PGS 147
MEAN 4469.73 SUM 657057
STD DEV 2055.63 VARIANCE 4225220
SKEWNESS -0.269257 KURTOSIS 0.553707
USS 355337911 CSS 61664084
CV 45.0895 STD MEAN 169.545
TIMEAN=0 FLD=ITI 0.0001
SCV PAW 5076.5 FPOD=ISI 0.0001
RUM = 0 142

QUANTILES(DEF=0)

100% MAX 9907 9907
75% Q3 5730 5730
50% MED 4557 4557
25% Q1 3492 3492
0% MIN 0 0
RANGE 9907
Q3-Q1 2240
MODE 0

EXTREMES

LOWEST TD HIGHEST TD
04005.03 80361002.03
04002.01 84191078.02
04018.01 96831073.03
04016.01 90041110.00
04015.01 99071142.01

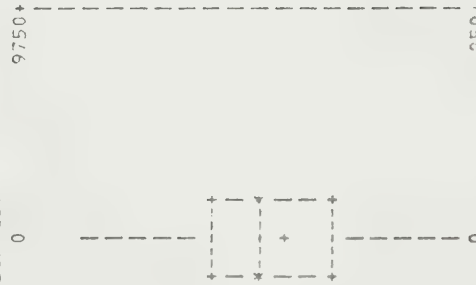
STEM LEAF

```

9 799
9
8 04
7 899
7 023
6 570999
6 111022222233
5 5555667770999
5 000011223333334
4 55667777770009
4 00000111111122223334444
3 5667778899
3 0112334
2 5666799
2 0223344
1 9
1
0
0 0000000011123

```

BOXPLOT



NORMAL PROBABILITY PLOT



FREQUENCY TABLE

VALUE	COUNT	PERCENTS	CELL	CUM
0	5	3.4	3.4	3.4
2	1	0.7	4.1	4.1
6	1	0.7	4.8	4.8
29	1	0.7	5.4	5.4
57	1	0.7	6.1	6.1
81	1	0.7	6.8	6.8
129	1	0.7	7.5	7.5
215	1	0.7	8.2	8.2
251	1	0.7	8.8	8.8
1910	1	0.7	9.5	9.5
2002	1	0.7	10.2	10.2
2166	1	0.7	10.9	10.9
2194	1	0.7	11.6	11.6

VALUE	COUNT	PERCENTS	CELL	CUM
3147	1	0.7	12.2	12.2
3213	1	0.7	12.9	12.9
3249	1	0.7	13.6	13.6
3249	1	0.7	14.3	14.3
3419	1	0.7	15.0	15.0
3422	1	0.7	15.6	15.6
3535	1	0.7	16.3	16.3
3613	1	0.7	17.0	17.0
3702	1	0.7	17.7	17.7
3728	1	0.7	18.4	18.4
3755	1	0.7	19.0	19.0
3777	1	0.7	19.7	19.7
3793	1	0.7	20.4	20.4
3793	1	0.7	21.1	21.1

1981 METEOROLOGICAL ATLAS SERIES
CMA: 835

13:08 WEDNESDAY, FEBRUARY 16, 1983

UNIVARIATE

VARIABLE=POPO_4

MOMENTS										EXTREMES			
N	140	SUM	WGTS	140	100% MAX	1255	99%	1255	LONEST	ID	HIGHEST	ID	
MEAN	367.429	SUM		51440	75% Q3	460	95%	998.75	01074.00		10801090.10		
STD DEV	275.905	VARIANCE		76123.6	50% MED	277.5	90%	768.5	01008.01		11151090.04		
SKEWNESS	1.36709	KURTOSIS		1.2474	25% Q1	177.5	10%	110.5	51001.01		1115110.00		
USS	29481700	CSS		10581174	0% MIN	0	5%	55.75	101076.04		12551076.03		
CV	75.0598	STD MEAN		23.3182	RANGE	1255	1%	0	201051.02		12551142.01		
12MEAN=0	15.7571	PROB> T		0.0001	Q3-Q1	282.5							
SGN RANK	4795.5	PROB> S		0.0001	MODE	200							
NUM	138												

MISSING VALUE
COUNT
% COUNT/NBBS

7
4.76

BOXPLOT

#

STEM LEAF

NORMAL PROBABILITY PLOT

1275+

0

2

12 55

1025+

0

2

11 11

775+

0

1

10 01

525+

0

2

9 57

275+

0

3

8 568

25+

0

1

7 57

0

0

2

10 01

0

0

3

9 57

0

0

3

8 568

0

0

3

7 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

0

3

9 57

0

0

3

10 01

0

</

UNIVARIATE

VARIABLE=POPO_4

FREQUENCY TABLE

VALUE	COUNT	PERCENTS CELL	CUM	VALUE	COUNT	PERCENTS CELL	CUM	VALUE	COUNT	PERCENTS CELL	CUM	VALUE	COUNT	PERCENTS CELL	CUM
0	2	1.4	1.4	190	1	0.7	26.4	310	1	0.7	57.9	610	1	0.7	82.9
5	1	0.7	2.1	195	1	0.7	27.1	315	3	2.1	60.0	640	1	0.7	83.6
10	1	0.7	2.9	200	6	4.3	31.4	325	1	0.7	60.7	650	1	0.7	84.3
20	2	1.4	4.3	205	1	0.7	32.1	330	2	1.4	62.1	660	1	0.7	85.0
55	1	0.7	5.0	210	1	0.7	32.9	335	2	1.4	63.6	675	1	0.7	85.7
70	2	1.4	6.4	215	4	2.9	35.7	345	1	0.7	64.3	725	1	0.7	86.4
75	1	0.7	7.1	220	2	1.4	37.1	355	1	0.7	65.0	730	3	2.1	88.6
85	1	0.7	7.9	225	4	2.9	40.0	380	1	0.7	65.7	735	1	0.7	89.3
95	1	0.7	8.6	230	2	1.4	41.4	385	2	1.4	67.1	755	1	0.7	90.0
110	2	1.4	10.0	240	1	0.7	42.1	410	1	0.7	67.9	770	1	0.7	90.7
115	1	0.7	10.7	245	1	0.7	42.9	420	1	0.7	68.6	815	1	0.7	91.4
125	1	0.7	11.4	250	1	0.7	43.6	425	1	0.7	69.3	855	1	0.7	92.1
130	2	1.4	12.9	255	3	2.1	45.7	430	2	1.4	70.7	860	1	0.7	92.9
135	1	0.7	13.6	265	2	1.4	47.1	450	1	0.7	71.4	885	1	0.7	93.6
140	1	0.7	14.3	270	3	2.1	49.3	460	6	4.3	75.7	950	1	0.7	94.3
150	1	0.7	15.0	275	1	0.7	50.7	465	1	0.7	76.4	975	1	0.7	95.0
155	3	2.1	17.1	280	1	0.7	51.4	470	1	0.7	77.1	1000	1	0.7	95.7
160	2	1.4	18.6	285	1	0.7	52.1	535	2	1.4	78.6	1010	1	0.7	96.4
165	4	2.9	21.4	290	1	0.7	52.9	565	2	1.4	80.0	1080	1	0.7	97.1
170	1	0.7	22.1	295	3	2.1	55.0	590	1	0.7	80.7	1115	1	0.7	97.8
175	4	2.9	25.0	300	2	1.4	56.4	595	1	0.7	81.4	1255	2	1.4	99.2
185	1	0.7	25.7	305	1	0.7	57.1	600	1	0.7	82.1				100.0

UNIVARIATE

VARIABLE=PCPOP0_4

MOMENTS

QUANTILES(DEF=4)

EXTREMES

	140	SUM	WGT	100% MAX	99%	18.2561	LONEST	ID	HIGHEST
MEAN	7.29607	SUM	1021.53	75% Q3	95%	14.4949	0(078.00)	18.8076(1090.06)	
STD. DEV.	3.53832	VARIANCE	12.5197	50% MED	90%	12.5795	0(008.01)	15.2684(1075.06)	
SKEWNESS	0.774325	KURTOSIS	0.411439	25% Q1	10%	3.87448	1.37845(032.02)	15.4551(1090.10)	
CV	48.4923	CSS	1740.24	0% MIN	5%	2.48991	1.566(038.00)	17.6683(1090.05)	
TIMEANF0	24.4001	STD MEAN	0.00001	RANGE	1%	0	1.79764(022.00)	16.5255(1090.06)	
SGN RANK	4795.5	PRUB> T	0.0001	Q3-Q1					
NUM T= 0	138	PRUB> S	0.0001	MUDE					

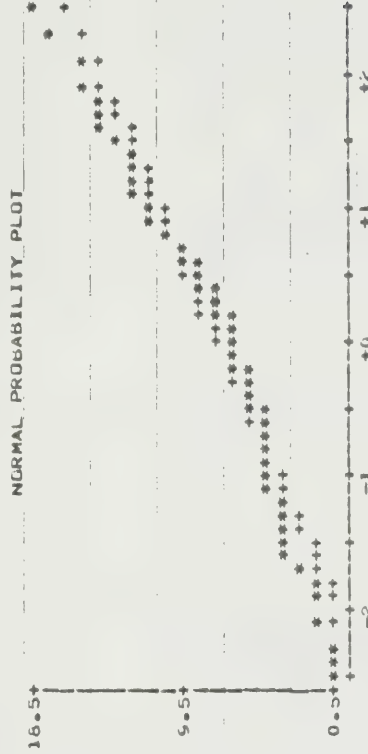
MISSING VALUE
COUNT
% COUNT/NOBS

7
4.76

BOXPLOT

STEM	LEAF	#
16	9	1
17	9	1
18	5	2
19	5	2
20	5	2
21	5	2
22	5	2
23	5	2
24	5	2
25	5	2
26	5	2
27	5	2
28	5	2
29	5	2
30	5	2
31	5	2
32	5	2
33	5	2
34	5	2
35	5	2
36	5	2
37	5	2
38	5	2
39	5	2
40	5	2
41	5	2
42	5	2
43	5	2
44	5	2
45	5	2
46	5	2
47	5	2
48	5	2
49	5	2
50	5	2
51	5	2
52	5	2
53	5	2
54	5	2
55	5	2
56	5	2
57	5	2
58	5	2
59	5	2
60	5	2
61	5	2
62	5	2
63	5	2
64	5	2
65	5	2
66	5	2
67	5	2
68	5	2
69	5	2
70	5	2
71	5	2
72	5	2
73	5	2
74	5	2
75	5	2
76	5	2
77	5	2
78	5	2
79	5	2
80	5	2
81	5	2
82	5	2
83	5	2
84	5	2
85	5	2
86	5	2
87	5	2
88	5	2
89	5	2
90	5	2
91	5	2
92	5	2
93	5	2
94	5	2
95	5	2
96	5	2
97	5	2
98	5	2
99	5	2
100	5	2

NORMAL PROBABILITY PLOT



FREQUENCY TABLE

VALUE	COUNT	PERCENTS	CELL	CUM	VALUE	COUNT	PERCENTS	CELL	CUM
0	2	1.4	1.4	1.4	4.09263	1	0.7	0.7	22.9
1.37845	1	0.7	2.1	2.1	4.09421	1	0.7	0.7	23.6
1.74764	1	0.7	2.9	2.9	4.09685	1	0.7	0.7	24.3
2.03308	1	0.7	3.6	3.6	4.79029	1	0.7	0.7	25.0
2.46212	1	0.7	4.3	4.3	4.81148	1	0.7	0.7	25.7
3.01787	1	0.7	5.0	5.0	4.85656	1	0.7	0.7	26.4
3.33434	1	0.7	5.7	5.7	4.8712	1	0.7	0.7	27.1
3.53821	1	0.7	6.4	6.4	4.89716	1	0.7	0.7	27.8
3.61905	1	0.7	7.1	7.1	4.95796	1	0.7	0.7	28.5
					4.97592	1	0.7	0.7	29.2

UNIVARIATE

VARIABLE=PCPUPT_4

FREQUENCY TABLE (CONT.)

VALUE	COUNT	PERCENTS CELL	CUM	VALUE	COUNT	PERCENTS CELL	CUM	VALUE	COUNT	PERCENTS CELL	CUM	VALUE	COUNT	PERCENTS CELL	CUM
4.99020	1	0.7	30.0	6.01018	1	0.7	47.9	11.2809	1	0.7	60.4	11.496	1	0.7	61.1
5.07737	1	0.7	30.7	6.02512	1	0.7	48.6	11.7119	1	0.7	61.8	11.879	1	0.7	62.5
5.10794	1	0.7	31.4	6.16761	1	0.7	49.3	12.0856	1	0.7	62.5	12.1257	1	0.7	63.2
5.12723	1	0.7	32.1	6.17024	1	0.7	50.0	12.1393	1	0.7	63.2	12.6284	1	0.7	63.9
5.26334	1	0.7	32.8	6.19472	1	0.7	50.7	12.6878	1	0.7	63.9	12.8386	1	0.7	64.6
5.29782	1	0.7	33.5	6.21699	1	0.7	51.4	13.0227	1	0.7	64.6	13.1175	1	0.7	65.3
5.32109	1	0.7	34.2	6.27233	1	0.7	52.1	13.1175	1	0.7	65.3	14.1175	1	0.7	66.0
5.32577	1	0.7	34.9	6.30025	1	0.7	52.8	14.5148	1	0.7	66.0	14.6313	1	0.7	66.7
5.35475	1	0.7	35.6	6.39562	1	0.7	53.5	14.6313	1	0.7	66.7	15.4604	1	0.7	67.4
5.40065	1	0.7	36.3	6.49627	1	0.7	54.2	15.4604	1	0.7	67.4	16.2604	1	0.7	68.1
5.43275	1	0.7	37.0	6.59226	1	0.7	54.9	16.2604	1	0.7	68.1	17.0551	1	0.7	68.8
5.44862	1	0.7	37.7	6.59233	1	0.7	55.6	17.0551	1	0.7	68.8	17.8503	1	0.7	69.5
5.59329	1	0.7	38.4	6.59233	1	0.7	56.3	17.8503	1	0.7	69.5	18.5255	1	0.7	70.2
5.69235	1	0.7	39.1	6.59233	1	0.7	57.0	18.5255	1	0.7	70.2				
5.79006	1	0.7	40.0	6.59233	1	0.7	57.7								
5.80515	1	0.7	40.7	6.59233	1	0.7	58.4								
5.85083	1	0.7	41.4	6.59233	1	0.7	59.1								
5.87034	1	0.7	42.1	6.59233	1	0.7	59.8								
5.90415	1	0.7	42.8	6.59233	1	0.7	60.5								
5.93643	1	0.7	43.5	6.59233	1	0.7	61.2								
5.97244	1	0.7	44.2	6.59233	1	0.7	61.9								
5.98115	1	0.7	44.9	6.59233	1	0.7	62.6								
5.98115	1	0.7	45.6	6.59233	1	0.7	63.3								
5.98115	1	0.7	46.3	6.59233	1	0.7	64.0								
5.98115	1	0.7	47.0	6.59233	1	0.7	64.7								
5.98115	1	0.7	47.7	6.59233	1	0.7	65.4								
5.98115	1	0.7	48.4	6.59233	1	0.7	66.1								
5.98115	1	0.7	49.1	6.59233	1	0.7	66.8								

3.3 GDMS

GDMS (Geographic Dot Mapping System) is a FORTRAN program developed at Statistics Canada which selects an optimum dot rate for the dot distribution map (see Figure 4). The actual size of the polygon representing each CT on the map is calculated. After inputting the number of occurrences in each CT for the particular theme (e.g., number of persons under five), GDMS calculates a density function, then sorts the CTs in order of density. Knowing the physical size of the dots that will be used on the map, it determines the minimum number of occurrences represented by each dot that will cause each CT to be completely covered by dots.

The rate calculated for the densest CT will produce dots that can be seen and counted everywhere on the map but there. Scanning down the list, lower density CTs become covered up as the dot rate declines. This is also shown graphically on the GDMS print-out (Figure 4).

Why would one want to use a lower dot rate if it meant covering up more and more CTs? The density function may be highly skewed, so that a large reduction in the dot rate might result in covering up only a few CTs while providing much more detail for the great majority. Apart from this, the subject-matter officer may wish to modify the dot rate either upwards or downwards to a "rounder" number (e.g., 50 instead of 55) or to a number used for another related theme.

The total number of dots across the CMA, which would be produced at each successive rate, will also be a factor in selecting an appropriate dot rate to use. Experience suggests that the maximum number of dots should be about 3,000 for the three small CMAs (St. John's, Halifax, Regina) and the six medium-sized CMAs, and 5,000 to 6,000 for Montréal, Toronto and Vancouver.

3.3 GDMS

Le GDMS (Geographic Dot Mapping System) est un programme FORTRAN élaboré à Statistique Canada, qui permet de choisir la densité optimale des points en vue de l'établissement de la carte de répartition par points (voir figure 4). La taille réelle du polygone représentant chaque SR sur la carte est calculée. Une fois introduit le nombre d'observations pour chaque SR relativement au thème choisi (par exemple, le nombre des personnes âgées de moins de 5 ans), le GDMS établit une fonction de densité et trie ensuite les SR par ordre de densité. Connaissant la grandeur réelle des points qui figureront sur la carte, le GDMS détermine le nombre minimal d'observations représentées par chaque point de façon que chaque SR soit complètement recouvert de points.

Selon le degré calculé pour le SR le plus dense, les points produits peuvent être vus et comptés partout ailleurs sur la carte, sauf dans ce secteur. Par balayage de la liste de haut en bas, les SR moins denses deviennent couverts de points à mesure que le degré de densité diminue. Ce phénomène est illustré graphiquement sur la sortie imprimée du GDMS (voir figure 4).

Pourquoi voudrait-on diminuer la densité des points si cela entraîne le recouvrement d'un plus grand nombre de SR? La fonction de densité peut être très asymétrique de sorte que toute diminution importante de la densité des points peut avoir pour effet de ne recouvrir que quelques SR et d'offrir par ailleurs un plus grand niveau de détail pour la plupart des autres SR. En outre, l'agent spécialisé peut décider d'augmenter ou de réduire la densité des points pour obtenir un chiffre "plus rond" (par exemple, 50 au lieu de 55) ou pour obtenir un nombre de points identiques à celui utilisé pour un thème connexe.

Un autre facteur dont on doit tenir compte au moment de déterminer la densité des points appropriée est le nombre total des points, pour l'ensemble de la RMR, produit selon chaque degré de densité. Par expérience, il semble que le nombre maximal de points devrait être limité à 3,000 pour les trois petites RMR (St. John's, Halifax et Regina) et six autres de taille moyenne et entre 5,000 et 6,000 pour Montréal, Toronto et Vancouver.

NAME OF VARIABLE: POP0_4
 EDMONTON DOT AREA: 514000
 TOTAL VALUE: 0.017

LINE NUMBER	CODE NUMBER	DENSITY (SQ. FT.)	DLIA DENSITY	VALUE	AREA (SQ. FT.)	UNROUNDED DOT RATING	EDMONTON TOTAL DOTS	LINE NUMBER
1	3762	54069.77	0.00	46500	0.86	914	5597	1
2	3763	47245.76	6824.01	111500	2.36	803	6406	2
3	3764	47058.82	186.94	64000	1.36	800	6430	3
4	3765	40951.22	107.60	38500	0.82	798	6446	4
5	3766	40972.22	5979.00	29500	0.72	687	7380	5
6	3767	40384.62	587.60	42000	1.04	687	7489	6
7	3768	39890.71	493.91	73000	1.83	678	7587	7
8	3769	39090.91	799.80	86000	2.20	645	7745	8
9	3770	35335.69	3755.22	100000	2.80	601	8259	9
10	3771	34008.51	667.16	125500	3.62	589	8733	10
11	3772	34375.00	293.51	16500	0.46	584	8608	11
12	3773	34353.74	21.26	101000	2.94	584	8608	12
13	3774	32634.11	1729.63	46000	1.41	584	9298	13
14	3775	30935.25	1688.96	43000	1.39	526	9779	14
15	3776	30872.48	62.77	46000	1.49	526	9798	15
16	3777	30555.36	316.92	27500	0.90	519	9931	16
17	3778	30454.55	101.01	33500	1.45	518	9931	17
18	3779	29591.84	862.71	72500	2.45	503	10227	18
19	3780	29109.59	482.25	42500	1.46	495	10392	19
20	3781	28125.00	984.39	85500	3.04	478	10762	20
21	3782	27232.14	892.86	61000	2.24	462	11110	21
22	3783	27155.17	76.97	31500	1.16	462	11134	22
23	3784	27020.20	134.97	53500	0.93	459	11207	23
24	3785	25806.45	1213.75	24000	0.93	439	11718	24
25	3786	25581.40	225.05	11000	0.43	435	11825	25
26	3787	24725.27	156.30	67500	2.73	420	12248	26
27	3788	24564.97	856.13	34500	1.16	418	12306	27
28	3789	24125.37	443.10	34500	1.43	398	12546	28
29	3790	23457.50	688.37	60000	2.56	397	12925	29
30	3791	23325.36	112.14	97500	4.18	392	12927	30
31	3792	23076.92	248.44	30000	1.30	389	13122	31
32	3793	22887.32	189.80	65000	2.84	389	13224	32
33	3794	22842.64	44.68	45000	1.97	388	13258	33
34	3795	22801.64	41.00	111500	4.89	388	13258	34
35	3796	22758.62	43.02	33000	1.45	387	13292	35
36	3797	22330.10	428.52	23000	1.03	380	13537	36
37	3798	21938.78	391.32	43000	1.90	373	13791	37
38	3799	21899.22	39.56	56500	2.58	372	13828	38
39	3800	21794.87	104.35	25500	1.17	371	13865	39
40	3801	21678.32	116.55	31000	1.43	369	13940	40
41	3802	21568.63	109.69	77000	3.57	367	14010	41
42	3803	20564.52	1004.11	25500	1.24	350	14697	42
43	3804	20564.52	0.90	25500	1.24	350	14697	43
44	3805	20229.91	335.51	26500	1.31	344	14953	44
45	3806	20175.44	53.57	46000	2.24	343	14997	45
46	3807	19871.79	303.65	15500	2.08	338	15219	46
47	3808	19711.54	160.25	41000	2.08	338	15355	47
48	3809	19613.26	98.28	35500	1.81	333	15447	48
49	3810	19402.99	210.27	13000	0.87	330	15586	49
50	3811	18664.05	738.94	95000	5.09	317	16227	50
51	3812	18595.04	69.01	22500	1.21	316	16278	51

GDMS

FIGURE 4

25

NAME OF VARIABLE TOTAL VALUE POPU_4
EDMONTON 5144000
DOT AREA 0.017

1	5	10	15	20	25	30	35	40	45	50	55	60	65	70	75	80	85	90	95	100	105	110	115	120
54069.77	X																							
53049.05	*																							
52024.32	*																							
51007.59	*																							
49986.86	*																							
48966.13	*																							
47945.40	*																							
46924.67	*	XXX																						
45903.94	*																							
44863.22	*																							
43862.49	*																							
42841.76	*																							
41821.03	*																							
40800.30	*	X																						
39779.57	*	XX																						
38758.84	*																							
37738.11	*																							
36717.39	*																							
35696.66	*																							
34675.93	*																							
33655.20	*																							
32634.47	*																							
31613.74	*																							
30593.01	*																							
29572.28	*																							
28551.56	*																							
27530.83	*																							
26510.10	*																							
25489.37	*																							
24468.64	*																							
23447.91	*																							
22427.18	*																							
21406.45	*																							
20385.73	*																							
19365.00	*																							
18344.27	*																							
17323.34	*																							
16302.81	*																							
15282.08	*																							
14261.35	*																							
13240.62	*																							
12219.90	*																							
11199.17	*																							
10178.44	*																							
9157.71	*																							
8136.98	*																							
7116.25	*																							
6095.32	*																							
5074.79	*																							
4054.05	*																							

LINE NUMBER

3.4 FASTCLUS

FASTCLUS is a SAS procedure which groups CTs into a small number of clusters based on similarities in their values on the derived variable (e.g., percentage of persons under 5).

FASTCLUS performs a disjoint cluster analysis. Given the desired number of clusters, it makes an initial, arbitrary selection of particular CTs as "class leaders". As it proceeds through the data set, the other CTs are either linked to or replace the existing class leaders based on their proximities in a Euclidean distance matrix.

Since FASTCLUS will produce different results depending on the order in which the clustering proceeds, its output alone cannot form a reliable basis for the selection of optimum class intervals. Cluster allocations are subsequently modified by a computer discriminant analysis procedure (Item 3.5) and the results compared against those which would be achieved by applying various standard distribution patterns to the data (Item 3.6).

This suite of procedures produces "optimum" class intervals for choropleth mapping. It is run three times to produce optimum intervals for a 3-class, a 4-class, and a 5-class map. The subject-matter officer selects which of the three he/she considers most appropriate for that particular theme and CMA, and then specifies the actual class intervals to be used on the map.

The most useful part of the FASTCLUS print-out (see Figure 5) is the list of CTs sorted in ascending order on the derived variable and showing the cluster allocations. Subject-matter officers are encouraged to use the sorted list of CTs as a sketch pad, to indicate the groupings achieved by the discriminant analysis technique and to check out any further modifications they may wish to make, when deciding upon the actual class intervals to use on the map.

3.4 FASTCLUS

Le FASTCLUS est un programme du SAS qui regroupe les SR en un petit nombre de grappes en fonction de correspondances entre les valeurs de la variable dérivée (par exemple, pourcentage des personnes âgées de moins de 5 ans).

FASTCLUS exécute une analyse désordonnée de grappes. Selon le nombre de grappes requis, il procède à une première sélection au hasard de certains SR désignés "guides de classe". À mesure que se poursuit le traitement de la série des données, les autres SR sont rattachés à ces guides de classe ou ils les remplacent selon leur proximité sur une matrice de distance euclidienne.

Étant donné que le FASTCLUS donnera des résultats différents selon l'ordre dans lequel s'effectue la mise en grappes, on ne peut baser le choix des intervalles de classe optimaux uniquement sur les résultats d'un tel traitement. Les répartitions par grappes sont par la suite modifiées grâce à l'application d'un programme informatique d'analyse discriminant (voir point 3.5) dont on compare les résultats avec ceux que l'on obtient par l'application de diverses méthodes courantes de distribution des données (voir point 3.6).

Cette suite de traitements permet de déterminer les intervalles de classe "optimaux" nécessaires à l'établissement des cartes choroplèthes. Le programme est passé trois fois pour obtenir les intervalles optimaux des cartes dont les observations peuvent porter sur trois, quatre ou cinq classes. L'agent spécialisé choisit celle des trois possibilités qu'il juge la plus appropriée au thème et à la RMR à l'étude, et détermine en conséquence les intervalles de classe réels qui serviront à la confection de la carte.

La partie la plus utile de la sortie imprimée du FASTCLUS (voir figure 5) est la liste des SR triés par ordre croissant d'importance de la variable dérivée et qui contient les répartitions par grappe. On conseille aux agents spécialisés d'inscrire directement sur cette liste des SR triés les groupements obtenus à l'aide de l'analyse discriminante et d'y préciser toute autre modification qu'ils souhaitent apporter, avant de déterminer les intervalles de classe devant être utilisés sur la carte.

1981 METROPOLITAN ATLAS SERIES
 CMA: 835 THEME: 3
 VAR: PCPOP0_4 MAXCLUSTERS 3

13:08 WEDNESDAY, FEBRUARY 10, 1983

PROC FASCLUS

CENTER MAXCLUSTERS = 3 ←

LEADER REPLACEMENTS: FIRST LEVEL = 2 SECOND LEVEL = 133

CLUSTER SUMMARY

CLUSTER	MEMBERS	MAX DISTANCE FROM CENTER
1	89	5.046615
2	7	3.682179
3	44	3.409971

CLUSTER CENTERS

CLUSTER	PCPOP0_4
1	5.046615022
2	16.196947162
3	10.707527466

DISTANCES BETWEEN CLUSTER CENTERS

CLUSTER	1	2	3
1	13.150332141	13.150332141	5.060912445
2	5.060912445	7.489419696	7.489419696
3			

FIGURE 5

1901 METROPOLITAN ATLAS SCRIPTS
CMA: 835 THEME: 3
VAR: PCPOPO_4 MAXCLUSTERS 4

13:08 WEDNESDAY, FEBRUARY 16, 1983 16

PROC FASCLUS

CENTER MAXCLUSTERS = 4 ←

LEADER REPLACEMENTS: FIRST LEVEL = 5 SECOND LEVEL = 129

CLUSTER SUMMARY

CLUSTER	MEMBERS	MAX DISTANCE FROM CENTER
1	82	4.609965
2	12	2.307014
3	2	0.3266086
4	44	2.733582

CLUSTER CENTERS

CLUSTER	PCPOPO_4
1	4.609964838
2	14.935439000
3	18.196947162
4	9.405721115

DISTANCES BETWEEN CLUSTER CENTERS

CLUSTER	1	2	3	4
1		10.325474162	13.586982324	4.795756277
2	10.525474162		3.261508162	5.529717885
3	13.586982324	3.261508162		8.791226047
4	4.795756277	5.529717885	8.791226047	

1981 METROPOLITAN ATLAS SERIES
CMA: 835 THEME: 3
VAR: PCPOP0_4 MAXCLUSTERS 5

PROC FACCLUS

CENTER MAXCLUSTERS = 5 ←
LEADER REPLACEMENTS: FIRST LEVEL = 4 SECOND LEVEL = 129

CLUSTER SUMMARY

CLUSTER MEMBERS MAX DISTANCE FROM CENTER

CLUSTER	MEMBERS	MAX DISTANCE FROM CENTER
1	6	2.039064
2	2	0.3286086
3	56	1.790114
4	53	2.638272
5	23	2.344661

CLUSTER CENTERS

CLUSTER	PCPOP0_4
1	0.000000000
2	18.196947162
3	4.248312838
4	7.975479532
5	13.402512408

DISTANCES BETWEEN CLUSTER CENTERS

CLUSTER	1	2	3	4	5
1	18.196947162	4.248312838	7.975479532	13.402512408	
2	4.248312838	13.948634325	10.221467030	4.794434754	
3	7.975479532	3.727106094	3.727166694	5.427032876	
4	13.402512408	9.154199570	5.427032876	13.402512408	
5					4.794434754

THEME: 3
 MAXCLUSTERS 5

←

UBS	CTNUM	CTCODE	PCPO0_4	CLUSTER
1	006.01	3760	0.00000	1
2	074.00	2215	0.00000	1
3	032.02	3775	1.37845	1
4	030.00	2169	1.56600	1
5	022.00	2161	1.79764	1
6	033.00	2172	2.03906	1
7	023.00	2162	2.46212	1
8	010.00	2149	3.01787	3
9	032.01	3774	3.33434	3
10	075.00	2211	3.51821	3
11	037.00	2176	3.61905	3
12	009.00	2148	3.67765	3
13	031.00	2170	3.70652	3
14	045.00	2184	3.87420	3
15	038.00	2177	3.87694	3
16	001.04	3755	3.92670	3
17	013.00	2152	3.99709	3
18	019.02	3773	4.00364	3
19	069.00	2208	4.13757	3
20	040.00	2179	4.15565	3
21	021.00	2160	4.18719	3
22	034.00	2173	4.21252	3
23	012.01	2761	4.34256	3
24	062.00	2201	4.38853	3
25	004.00	2143	4.41104	3
26	005.01	3758	4.45736	3
27	035.00	2174	4.49871	3
28	016.02	3771	4.49944	3
29	011.00	2150	4.51427	3
30	008.00	2147	4.52182	3
31	068.00	2207	4.67290	3
32	042.02	3777	4.69263	3
33	017.00	2156	4.69421	3
34	041.00	2180	4.69885	3
35	046.00	2185	4.79029	3
36	020.00	2159	4.81148	3
37	053.00	2192	4.85656	3
38	039.00	2176	4.87120	3
39	014.00	2153	4.89716	3
40	047.00	2186	4.95798	3
41	043.00	2182	4.97592	3
42	034.00	2193	4.99029	3
43	029.00	2168	5.07737	3
44	111.00	2772	5.07849	3
45	002.01	2756	5.10794	3
46	057.00	2196	5.12723	3
47	036.00	2175	5.29234	3
48	025.00	2164	5.29782	3
49	049.00	2188	5.32167	3
50	026.00	2165	5.32577	3
51	048.00	2187	5.35475	3
52	024.00	2163	5.40065	3
53	052.02	3781	5.43275	3
54	076.01	3794	5.44662	3

discrim 37.

OBS	CTNUM	CTCODE	PCPO0_4	CLUSTER
55	066.00	2205	5.59329	3
56	015.02	3769	5.69235	3
57	051.01	3778	5.79006	3
58	056.00	2195	5.80515	3
59	063.00	2202	5.85683	3
60	059.00	2198	5.87054	3
61	012.02	2762	5.90418	3
62	005.00	2142	6.03843	3
63	001.01	3752	6.17284	3
64	002.02	2759	6.25153	3
65	073.00	2212	6.26118	3
66	064.01	3784	6.35593	3
67	071.00	2210	6.39439	3
68	067.00	2206	6.48962	3
69	075.03	3789	6.49351	3
70	055.00	2194	6.51196	3
71	077.01	3796	6.53071	3
72	027.00	2166	6.57473	3
73	065.02	3787	6.60189	3
74	007.01	3766	6.70241	3
75	065.01	3786	6.73744	3
76	061.00	2200	6.85116	3
77	070.00	2209	6.85643	3
78	060.01	3782	6.87285	3
79	120.01	3817	6.87285	3
80	04.00	2183	6.89599	3
81	007.02	3767	6.91077	3
82	100.00	3011	6.94716	3
83	005.02	3759	7.01262	3
84	075.01	2763	7.24002	3
85	050.00	2189	7.37179	3
86	121.03	3823	7.62631	3
87	060.02	3783	7.67357	3
88	076.04	3801	7.75194	3
89	076.02	3795	7.86106	3
90	121.02	3822	7.88644	3
91	106.00	2770	7.89908	3
92	051.02	3779	7.96813	3
93	101.01	3813	8.01018	3
94	105.01	3815	8.02512	3
95	120.02	3818	8.16761	3
96	001.03	3754	8.17025	3
97	001.05	3756	8.19472	3
98	028.00	2167	8.21699	3
99	104.00	2768	8.27233	3
100	075.02	2764	8.30025	3
101	002.03	3757	8.39970	3
102	120.03	3819	8.98562	3
103	140.00	2775	9.19627	3
104	078.01	3798	9.25926	3
105	018.00	2157	9.30233	3
106	103.00	2767	9.32118	3
107	058.00	2197	9.57008	3
108	001.02	3753	9.70478	3

discrim 6.3%

1981 METROPOLITAN ATLAS SERIES
CMA: 835
VAR: PCPOPO_4 MAXCLUSTERS 5

ONS	CINUM	CICODE	PCPOPO_4	CLUSTER
109	042.01	3776	10.0065	4
110	141.00	2776	10.2912	4
111	077.02	3797	10.4031	4
112	102.00	2766	10.4614	4
113	142.02	3625	10.4877	4
114	105.02	3616	10.4982	4
115	006.05	3764	10.4138	4
116	096.02	3804	11.0585	5
117	006.03	3762	11.1484	5
118	110.00	2771	11.2509	5
119	090.03	3805	11.4960	5
120	090.01	3603	11.7119	5
121	078.05	3602	11.8250	5
122	078.02	3799	11.8779	5
123	006.02	3765	12.0589	5
124	006.02	3761	12.0858	5
125	121.01	3821	12.1257	5
126	006.04	3763	12.1393	5
127	075.04	3796	12.6284	5
128	142.01	3624	12.6678	5
129	120.04	3820	12.8386	5
130	078.03	3800	12.9609	5
131	101.02	3814	13.0227	5
132	075.05	3791	13.1175	5
133	090.04	3806	14.1175	5
134	090.09	3611	14.5148	5
135	075.07	3793	14.6313	5
136	090.08	3810	14.8076	5
137	075.06	3792	15.2684	5
138	090.10	3612	15.4551	5
139	090.05	3807	17.8683	2
140	090.06	3808	18.5256	2

discrim 10.38

16.2

3.5 DISCRIM

DISCRIM is a SAS procedure which performs discriminant analysis on the sorted FASTCLUS output. DISCRIM examines the existing CT clusters and, through a process which minimizes the within-group variances and maximizes the between-group variance, is able to re-allocate individual CTs from one group to another in order to achieve a better clustering of CTs.

The most useful part of the DISCRIM print-out (see Figure 6) is the listing of those CTs which have been re-allocated by the procedure. By marking these on the sorted FASTCLUS print-out, one is able to determine the revised class "breaks" which the computer analysis suggests would be most appropriate for that theme and CMA.

3.5 DISCRIM

Le DISCRIM est un programme du SAS qui exécute une analyse discriminante à partir des résultats du tri du programme FASTCLUS. Le DISCRIM permet de faire l'examen des grappes de SR grâce à un processus qui minimise la variation à l'intérieur d'un groupe donné et maximise la variation entre les groupes; les différents SR peuvent être répartis à nouveau d'un groupe à un autre de sorte qu'on obtient des grappes plus adéquates.

La partie la plus utile de la sortie imprimée du DISCRIM (voir figure 6) est la liste de ces SR qui ont fait l'objet d'une nouvelle répartition à la suite du traitement. En les marquant sur la sortie imprimée du FASTCLUS, il est possible de modifier en conséquence les intervalles de classe du thème et de la RMR à l'étude en fonction des résultats de cette analyse par ordinateur.

1981 METROPOLITAN ATLAS SERIES
 CMA: 835
 VAR: PCPOP0_4 MAXCLUSTERS 3

13:08 WEDNESDAY, FEBRUARY 16, 1983

DISCRIMINANT ANALYSIS

CLUSTER	FREQUENCY	PRIOR PROBABILITY
1	59	0.3333333
2	7	0.3333333
3	44	0.3333333
TOTAL	110	1.0000000

} 3 classes

FIGURE 6

DISCRIMINANT ANALYSIS

GENERALIZED SQUARED DISTANCE FUNCTION:

$$D^2(x) = (x - \bar{x}_j)' \cdot \text{COV}^{-1}(x - \bar{x}_j)$$

CLASSIFICATION RESULTS FOR CALIBRATION DATA: WORK.MAXC3
 POSTERIOR PROBABILITY OF MEMBERSHIP IN EACH CLUSTER:

$$\text{PR}(J|X) = \frac{\exp(-0.5 D_K^2(x))}{\sum_K \exp(-0.5 D_K^2(x))}$$

POSTERIOR PROBABILITY OF MEMBERSHIP IN CLUSTER:

UBS	FROM CLUSTER	CLASSIFIED INTO CLUSTER	1	2	3
88	1	3 *	0.4916	0.0000	0.5084
89	1	3 *	0.4412	0.0000	0.5588
132	3	2 *	0.0000	0.5021	0.4978
133	3	2 *	0.0000	0.0762	0.1238

* MISCLASSIFIED OBSERVATION

(3 classes)

1901 MEIKUPULITAN ATLAS SERIES
CMA: 835
THEME: 3
VAR: PCPOPO_4 MAXCLUSTERS 4

13:08 WEDNESDAY, FEBRUARY 10, 1983 20

DISCRIMINANT ANALYSIS

CLUSTER	FREQUENCY	PRIOR PROBABILITY
1	62	0.25000000
2	12	0.25000000
3	2	0.25000000
4	44	0.25000000
TOTAL	140	1.00000000

} 4 classes

1981 METROPOLITAN ATLAS SERIES
CMA: 835
VAR: PCPOPO_4 MAXCLUSTERS 4

13:00 WEDNESDAY, FEBRUARY 10, 1983

25

DISCRIMINANT ANALYSIS

CLASSIFICATION RESULTS FOR CALIBRATION DATA: WORK-MAXC4

GENERALIZED SQUARED DISTANCE FUNCTION:

POSTERIOR PROBABILITY OF MEMBERSHIP IN EACH CLUSTER:

$$D_j^2(x) = (x - \bar{x}_j)^T \text{COV}_j^{-1} (x - \bar{x}_j)$$

$$PR(j|x) = \exp(-0.5 D_j^2(x)) / \sum_K \exp(-0.5 D_K^2(x))$$

POSTERIOR PROBABILITY OF MEMBERSHIP IN CLUSTERS:

		FROM CLUSTER	CLASSIFIED INTO CLUSTER	1	2	3	4
035							
83			1 *	0.5879	0.0001	0.0000	0.4120
120		4	2 *	0.0001	0.5263	0.0001	0.4735
121		4	2 *	0.0000	0.5796	0.0002	0.4202
122		4	2 *	0.0000	0.6039	0.0002	0.3956
123		4	2 *	0.0000	0.6827	0.0004	0.3169
124		4	2 *	0.0000	0.6937	0.0004	0.3059
125		4	2 *	0.0000	0.7096	0.0004	0.2900
126		4	2 *	0.0000	0.7149	0.0004	0.2846

* MISCLASSIFIED OBSERVATION

(4 classes)

DISCRIMINANT ANALYSIS

CLUSTER	FREQUENCY	PRIOR PROBABILITY
1	6	0.20000000
2	2	0.20000000
3	50	0.20000000
4	55	0.20000000
5	23	0.20000000
TOTAL	146	1.00000000

5 classes

1981 METROPOLITAN ATLAS SERIES
CMA: 835
THEME: 3
VAR: PCPOP0_4 MAXCLUSTERS 5

DISCRIMINANT ANALYSIS

CLASSIFICATION RESULTS FOR CALIBRATION DATA: WORK.MAXCS

GENERALIZED SQUARED DISTANCE FUNCTION: POSTERIOR PROBABILITY OF MEMBERSHIP IN EACH CLUSTER:

$$D_j^2(x) = (x - \bar{x}_j)' \text{COV}_j^{-1} (x - \bar{x}_j) \quad \text{PR}(j|x) = \exp(-.5 D_j^2(x)) / \sum_K \exp(-.5 D_K^2(x))$$

POSTERIOR PROBABILITY OF MEMBERSHIP IN CLUSTER:

OBS	FROM CLUSTER	CLASSIFIED INTO CLUSTER	1	2	3	4	5
7	3	1 *	0.7783	0.0000	0.2217	0.0000	0.0000
63	4	3 *	0.0001	0.0000	0.5066	0.4111	0.0000
64	4	3 *	0.0001	0.0000	0.5401	0.4598	0.0000
65	4	3 *	0.0000	0.0000	0.5341	0.4659	0.0000
111	4	5 *	0.0000	0.0000	0.0000	0.4743	0.5257
112	4	5 *	0.0000	0.0000	0.0000	0.4011	0.5989
113	4	5 *	0.0000	0.0000	0.6000	0.3953	0.0047
114	4	5 *	0.0000	0.0000	0.0000	0.3858	0.6142
115	4	5 *	0.0000	0.0000	0.0000	0.2881	0.7119

* MISCLASSIFIED OBSERVATION

(5 classes)

3.6 CHANG

CHANG is a FORTRAN program developed by Professor K. Chang at the University of North Dakota. CHANG computes class intervals for the data by applying seven standard classing methods. These standard methods are: (1) "equal frequency", creating classes with an equal number of CTs in each; (2) "equal step", for classes of equal width; (3) "arithmetic", for classes of linearly increasing width; (4) "geometric", for classes of exponentially increasing width; (5) "reciprocal", for classes of decreasing width; (6) "mean and standard deviation", for classes of equal frequency, assuming a normal distribution; (7) "nested means", for classes hierarchically divided using means. In addition to the above, the production team keys in the "optimum" class intervals as determined by the discriminant analysis procedure.

CHANG computes a "D value" for each method, which is the sum of the within-group variances it produces. Lower D values indicate a better fit to the data. The subject-matter officer can use the D value to compare the DISCRIM results with those produced by the various standard methods.

The CHANG program is also run three times to produce results which can be applied to a 3-class, 4-class, or 5-class choropleth map.

The CHANG analysis (see Figure 7), together with the summary statistics from UNIVARIATE and the sorted list of CTs from FASTCLUS, provides the basis for the subject-matter officer's determination of both the appropriate number of classes and the appropriate class intervals for the choropleth map.

3.6 CHANG

Le CHANG est un programme FORTRAN mis au point par le professeur K. Chang de l'"University of North Dakota". Le CHANG permet de calculer les intervalles de classe appropriés à l'aide de sept méthodes-types de classification des données. En voici la liste: (1) "effectif égal", classes comprenant un nombre égal de SR; (2) "intervalle égal", pour les classes d'amplitude égale; (3) "progression arithmétique", classes dont l'amplitude croît linéairement; (4) "progression géométrique", pour les classes dont l'amplitude croît de façon exponentielle; (5) "réciproque", pour les classes d'amplitude décroissante (6) "écart moyen et écart-type", pour les classes d'effectif égal obéissant à une loi de distribution normale; (7) "moyennes emboîtées", pour les classes réparties hiérarchiquement en fonction de moyennes. En plus des résultats obtenus à la suite de l'application de ces méthodes, l'équipe de production introduit les intervalles de classe "optimaux" déterminés par l'analyse discriminante.

Le programme CHANG calcule une "valeur D" qui correspond à la somme des variances à l'intérieur des groupes obtenues avec chacune de ces méthodes. Les valeurs D plus faibles témoignent d'un meilleur ajustement des données. L'agent spécialisé peut se servir de la valeur D pour comparer ces résultats avec ceux de l'analyse effectuée par le DISCRIM.

Le programme CHANG est lui aussi passé trois fois pour produire des résultats pouvant être appliqués à l'établissement des cartes choroplèthes de trois, quatre ou cinq classes.

Pour déterminer le nombre approprié de classes ainsi que les intervalles de classe convenant à l'établissement des cartes choroplèthes, l'agent spécialisé doit donc se fonder sur l'analyse CHANG (voir figure 7) de même que sur les statistiques sommaires fournies par le programme UNIVARIATE et la liste des SR triés établie par le programme FASTCLUS.

→ EQUAL FREQUENCY METHOD RESULTS*****

CLASS BREAKS ARE:

4.50
5.80
7.00
10.50

COMPIATION OF CLASS FREQUENCIES

CLASS	FREQUENCY
1	28
2	24
3	28
4	21
5	21

} 5 CLASSES

COMPUTATION OF U VALUES

CLASS	U VALUE
1	0.2258773
2	0.0000001
3	0.0031860
4	0.0084114
5	0.013897

→ TOTAL U VALUE: 0.298866540 ←

→ EQUAL STEP METHOD RESULTS*****

CLASS BOUNDS ARE:

4.82
8.21
11.66
15.06

COMPIATION OF CLASS FREQUENCIES:

CLASS	FREQUENCY
1	34
2	62
3	21
4	17
5	4

COMPUTATION OF D VALUES

CLASS	D VALUE
1	0.23324
2	0.01010
3	0.00124
4	0.00968
5	0.00012

TOTAL D VALUE 0.25438 ←

→ ARITHMETIC METHOD RESULTS*****

CLASS BREAKS ARE:

2.34
4.02
6.24
12.80

COMPUTATION OF CLASS FREQUENCIES:

CLASS	FREQUENCY
1	5
2	29
3	62
4	31
5	11

COMPUTATION OF J VALUES

CLASS	J VALUE
1	0.02083
2	0.07402
3	0.01010
4	0.09524
5	0.01740

TOTAL J VALUE 0.12759 ←

→ GEOMETRIC METHOD RESULTS*****

CLASS BREAKS ARE:

2.35
3.43
6.59
11.04

COMPUTATION OF CLASS FREQUENCIES:

CLASS	FREQUENCY
1	4
2	10
3	55
4	44
5	25

COMPUTATION OF U VALUES

CLASS	U VALUE
1	0.04152
2	0.05214
3	0.00616
4	0.03498
5	0.05002

TOTAL U VALUE 0.19192 ←

→ RECIPROCAL METHOD RESULTS*****

CLASS BREAKS ARE:

1.72
2.22
3.14
5.37

COMPIATION OF CLASS FREQUENCIES:

CLASS	FREQUENCY
1	2
2	2
3	2
4	42
5	90

COMPUTATION OF D VALUES

CLASS	D VALUE
1	0.00101
2	0.00934
3	0.00847
4	0.02944
5	0.34051

TOTAL D VALUE 0.38938 ←

→ MEAN AND STANDARD DEVIATION METHOD RESULTS**

CLASS BREAKS ARE:

4.51
6.54
8.28
10.30

COMPUTATION OF CLASS FREQUENCIES:

CLASS	FREQUENCY
1	29
2	41
3	27
4	12
5	30

COMPUTATION OF U VALUES

CLASS	U VALUE
1	0.22449
2	0.00101
3	0.00015
4	0.00151
5	0.07170

TOTAL U VALUE 0.29445

*****ERROR TYPE I*****

4. DATA RETRIEVAL

4.1 General Remarks

CT-level data for the metropolitan atlases are retrieved by the production team from the 2A (100% data) and 2B (20% sample data) Profile user summary tapes, according to specifications provided by subject-matter. CMA, province and Canada level comparative data are supplied directly by the subject-matter officer.

The atlases conform with the data suppression policy adopted for the CT Profile Series. Unless otherwise specified by the subject-matter officer, all CTs with a total 1981 population of less than 25 will be suppressed. Suppressed CTs are excluded from the analytical programs and identified by a special shading pattern on the maps and graphs.

4.2 Specification of Variables for Retrieval

The subject-matter officer identifies all of the input variables that will be required for his/her theme on a prepared form. The officer assigns each input variable an upper-case letter for identification purposes. Any operations which may be required to collapse input variables are indicated with a mathematical expression using the alphabetic identifiers.

To illustrate the above, take the theme "age group under five years" as an example.

The input variables are:

A = population, 1981
B = male, 0-4 years
C = female, 0-4 years

The collapsed variable is:

Q = population, 0-4 years

The collapsing operation is indicated as:

$Q = B + C$

Collapsing is performed on unrounded data, and the total is then randomly rounded.

4. EXTRACTION DES DONNÉES

4.1 Observations générales

Les données relatives aux SR devant servir à l'établissement des atlas métropolitains sont extraites, par l'équipe de production, des bandes sommaires des utilisateurs de la Série des profils 2A (données recueillies auprès de l'ensemble de la population) et 2B (données recueillies auprès d'un échantillon de 20 % de la population), selon les directives fournies par chaque agent spécialisé.

La politique d'élimination des données est la même que celle qui a été adoptée dans le cas de la Série des profils des SR. À moins d'un avis contraire de l'agent spécialisé tous les SR dont la population totale en 1981 était inférieure à 25 doivent être éliminés. Ceux-ci sont donc exclus des programmes d'analyse mais doivent être différenciés à l'aide d'un motif hachuré distinct sur les cartes et les graphiques.

4.2 Détermination des variables aux fins de l'extraction des données

L'agent spécialisé indique sur une formule conçue spécialement à cette fin toutes les variables d'entrée requises pour son thème. Chacune de ces variables est désignée à l'aide d'une lettre majuscule. S'il est nécessaire de combiner certaines variables d'entrée, l'opération est indiquée par une formule mathématique qui comprend les identificateurs alphabétiques.

Pour illustrer ce qui précède, prenons l'exemple du thème "groupe des personnes âgées de moins de 5 ans".

Les variables d'entrée sont:

A = population, 1981
B = sexe masculin, 0-4 ans
C = sexe féminin, 0-4 ans

La variable combinée est:

Q = population, 0-4 ans

L'opération de combinaison est indiquée de la manière suivante:

$Q = B + C$

L'opération de combinaison est exécutée à partir de données non arrondies et on procède ensuite à l'arrondissement aléatoire du total.

4.3 Specification of Variables for Mapping

All counts are randomly rounded at the time of retrieval, after any required collapsing operations have been performed, but before any percentages are calculated and the data passed to the analytical programs.

The subject-matter officer specifies the variables for mapping using the same alphabetic identifiers. Following on the same age group example, the dot distribution map would use the collapsed variable:

Q = population, 0-4 years

The choropleth map would use the derived variable:

V = percentage, 0-4 years

The mathematical operation is indicated as:

$$V = Q/A \times 100$$

Percentages are reported to one decimal place.

Provision has been made for calculations involving a constant, supplied by the subject-matter officer. Take the choropleth map for the theme "working in the CT" as an example.

The input variable is:

A = Total, working in the CT

The constant (supplied by subject-matter) is:

N = total CMA employed labour force

The derived variable is:

V = percentage, working in the CT

where:

$$V = A/N \times 100$$

The subject-matter officer is supplied with a print-out of all input variables (after collapsing, if applicable) and derived variables for the purpose of verifying the retrieval and the mathematical operations for his/her theme.

4.3 Détermination des variables aux fins de l'établissement des cartes

Tous les chiffres sont arrondis de manière aléatoire au moment de l'extraction des données, après l'exécution des opérations de combinaison (s'il y a lieu), mais avant le calcul de tout pourcentage et le traitement des données par les programmes d'analyse.

L'agent spécialisé désigne les variables qui devront servir à l'établissement des cartes, à l'aide des mêmes identificateurs alphabétiques. Si nous reprenons l'exemple du même groupe d'âge, la carte de répartition par points porterait sur la variable combinée:

Q = population, 0-4 ans

La carte choroplèthe porterait sur la variable dérivée:

V = pourcentage, 0-4 ans

La formule mathématique s'écrit :

$$V = Q/A \times 100$$

Les pourcentages sont arrondis à une seule décimale.

On a tenu compte du fait que, dans certains cas, les calculs doivent faire intervenir une constante, laquelle est fournie par l'agent spécialisé. Prenons à titre d'exemple une carte choroplèthe analysant le thème "travaillant dans le SR".

La variable d'entrée est:

A = total des personnes employées dans le SR

La constante (précisée par l'agent spécialisé) est:

N = population active totale occupée dans la RMR

La variable dérivée est:

V = pourcentage, travaillant dans le SR

où:

$$V = A/N \times 100$$

L'agent spécialisé se voit remettre une liste imprimée de toutes les variables d'entrée (après combinaison, s'il y a lieu) et des variables dérivées afin qu'il puisse vérifier l'exactitude des opérations d'extraction des données et de calculs mathématiques ayant trait à son thème.

5. OPERATIONAL PROCEDURES

5.1 General Remarks

The general work flow for the metropolitan atlas project is shown in Figure 8. Major subject-matter inputs are required at three stages: (1) specification of variables for retrieval and mapping; (2) preparation of text and comparative data; (3) selection of class intervals and dot rate for mapping. Forms are supplied for submitting the information required at each stage. After processing each of these inputs, the production team returns material (print-outs, preliminary plots, and the final print-ready maps) for verification purposes.

A single date is established for the specification of all variables to allow them all to be retrieved in a single operation. In the second stage, text and comparative data for all 12 CMAs are supplied by groups of themes. In the final stage, class intervals and dot rates for all themes are supplied by groups of CMAs. This three-stage process is first applied to the 2A themes, and later to the 2B themes, as their respective user summary tapes become available.

5. MÉTHODES OPÉRATIONNELLES

5.1 Observations générales

Le déroulement général des opérations du projet des atlas métropolitains est illustré à la figure 8. Les principales données d'entrée par domaine spécialisé doivent être fournies aux trois étapes suivantes: (1) détermination des variables aux fins de l'extraction des données et de l'établissement des cartes; (2) production du texte et des données comparatives; (3) détermination des intervalles de classe et de la densité des points aux fins de l'établissement des cartes. Des formules ont été conçues pour la présentation de l'information requise à chacune de ces étapes. Après le traitement de ces données d'entrée, l'équipe de production remet le matériel (listes imprimées, tracés préliminaires et cartes finales prêtes à être imprimées) à l'agent spécialisé pour qu'il le vérifie.

À la première étape, un seul délai doit être fixé de manière que l'extraction de toutes les variables se fasse en une seule opération. À la deuxième étape, les textes et les données comparatives servant à l'ensemble des 12 RMR sont fournis par groupes de thèmes. À la dernière étape, les intervalles de classe et la densité des points choisis pour l'ensemble des thèmes sont regroupés par RMR. Ce traitement en trois étapes est appliqué d'abord aux thèmes particuliers à la formule 2A du recensement et ensuite aux thèmes de la formule 2B, à mesure qu'on a accès aux bandes sommaires des utilisateurs de ces deux séries.

GENERAL
WORK
FLOW

1.

Specify Variables
(SM)
Sélection des variables
(AS)DÉROULEMENT
DES
ACTIVITÉS

2.

Prepare Text and
Comparative Data (SM)
Préparation du texte et des
données comparatives (AS)Key Enter Text
(PR)
Introduction du texte
(ÉP)(SM)
Text
OK ?
(ÉP)
Texte
OK ?
No / Non
Yes / Oui(SM)
Data
OK ?
(AS)
Données
OK ?
No / Non
Yes / OuiAnalyse Data
(PR)
Analyse des données
(ÉP)

3.

Specify Class Intervals
and Dot Rate (SM)
Sélection des intervalles de classe
et du débit des points (AS)Plot Maps
(PR)
Tracage des cartes
(ÉP)(SM)
Maps
OK ?
(AS)
Cartes
OK ?
No / Non
Yes / OuiPrint Maps
(PR)
Impression des cartes
(ÉP)SM - Subject Matter Officer
AS - Agent spécialiséPR - Production Team
EP - Équipe de production1/2/3 - Major Subject Matter Inputs
1/2/3 - Principaux paramètres

5.2 Specification of Variables

All variables required for retrieval and mapping of the theme, together with the mathematical operations to be performed on theme, are specified by the subject-matter officer on the Form 1 (Figure 9). Refer to Chapter 4 of this guide for an explanation of the requirements at this stage. If the data retrieval and/or mathematical operations for a given theme are not the same for all 12 CMAs, use a separate form for each.

After processing this input, the production team will provide print-outs, by CMA and CT, of the retrieved and derived variables to the subject-matter officer for verification. The production team will then run the data analysis programs and provide the outputs to the subject-matter officer for selection of class intervals and dot rate (see Item 5.4).

5.2 Détermination des variables

L'agent spécialisé inscrit sur la formule 1 (voir figure 9) tous les renseignements relatifs aux variables nécessaires à l'extraction des données et à l'établissement des cartes thématiques ainsi qu'aux calculs mathématiques à exécuter. Pour de plus amples explications sur les exigences particulières de cette étape, se reporter au chapitre 4 du présent guide. Si les directives touchant l'extraction des données ou les calculs mathématiques pour un thème donné ne sont pas les mêmes pour l'ensemble des 12 RMR, il convient d'utiliser une formule distincte pour chacune des régions.

Après le traitement de ces données d'entrée, l'équipe de production soumet à la vérification de l'agent spécialisé les listes imprimées des variables extraites et dérivées, réparties entre les RMR et les SR. L'équipe de production passe ensuite les programmes d'analyse des données et fait parvenir les résultats à l'agent spécialisé afin qu'il puisse déterminer les intervalles de classe et la densité des points (voir point 5.4).

Metropolitan Atlas Series

FORM 1

Specification of Variables for Retrieval and Mapping (see Guide item 4)

Theme No: 3 Theme Name: POPULATION UNDER 5 YEARS Specify data base: ☒ 2A ☐ 2B All CMA's: ☒ or specify CMA: _____

1. Identify all input variables from UST list (see Guide item 4.2)

alphanumeric identifier	Variable name	Variable description
A = 3	_____	POPULATION, 1981
B = 7	_____	MALE, 0-4 YEARS
C = 19	_____	FEMALE, 0-4 YEARS
D =	_____	_____
E =	_____	_____
F =	_____	_____
G =	_____	_____
H =	_____	_____
I =	_____	_____
J =	_____	_____
K =	_____	_____
L =	_____	_____
M =	_____	_____

3. Collapsed variables (see Guide item 4.2)

alphanumeric identifier	Variable name (assigned by GCG)	Variable description	Specify operations to collapse variables
Q =	_____	POPULATION 0-4 YEARS	where Q = B + C
R =	_____	_____	where R =
S =	_____	_____	where S =
T =	_____	_____	where T =
U =	_____	_____	where U =

4. Derived variables (see Guide item 4.3)

alphanumeric identifier	Variable name (assigned by GCG)	Variable description	Specify operations to derive variables
V =	_____	PERCENT 0-4 YEARS	where V = $Q/A \times 100$
W =	_____	_____	where W =
X =	_____	_____	where X =
Y =	_____	_____	where Y =
Z =	_____	_____	where Z =

2. Specify constants, if applicable (see Guide item 4.3)

alphanumeric identifier	Constant value	Description
N =	_____	_____
O =	_____	_____
P =	_____	_____

5. Dot map shows variable: Q6. Choropleth map shows variable: V

7. Special instructions re CT exclusions (CTs with total population (25 exclude unless otherwise specified):

☐ No CT exclusions; or

☐ Exclude CT if variable _____ less than _____

8. Send output listings to:

name: _____
location: _____
telephone: _____

For use by subject matter officer(s)

- Form completed by: _____ date: _____
- Form approved and forwarded to GCG by: _____ date: _____
- Printouts received and variables verified by: _____ date: _____
- Variables rejected-see memo attached: _____ date: _____
- Variables accepted-GCG notified by: _____ date: _____

For use by GCG production team:

- Form received and variable names assigned by: _____ date: _____
- Retrieval job coded and submitted by: _____ date: _____
- Printouts checked and forwarded to subject matter by: _____ date: _____
- Variables rejected-corrective action taken by: _____ date: _____
- Variables accepted-files stored & documented by: _____ date: _____

5.3 Preparation of Text and Comparative Data

All textual material and comparative data required for the presentation of the theme are supplied by the subject-matter officer on Forms 2 and 3 (Figures 10 and 11). Refer to Chapter 2 of this guide for an explanation of the requirements at this stage.

Form 2 is for text. Submit two copies of this form for each theme - one for the choropleth map page and one for the dot map page. Note that all text must be bilingual and must conform to the space limitations indicated on the form. Set up the punctuation, spacing and use of upper and lower case letters exactly as you wish them to appear. If the text for a given theme is not the same for all 12 CMAs, use a separate pair of forms for each.

Form 3 is used to supply comparative CMA/province/Canada data. For the cumulative bar graph (if applicable) submit comparative data for up to five major variable classifications, plus totals, as rounded counts in Rows 1 to 5 and T. For the comparative data table, submit percentages calculated from rounded data (or other appropriate values, e.g., average incomes) in Row P.

After processing this input, the production team will provide preliminary plots, with text and comparative data only, to the subject-matter officer for verification.

5.3 Préparation du texte et des données comparatives

L'agent spécialisé utilise les formules 2 et 3 (voir figures 10 et 11) pour préciser le texte et les données comparatives qui sont nécessaires à la présentation du thème choisi. Se reporter au chapitre 2 du présent guide pour une explication des exigences particulières de cette étape.

La formule 2 est réservée au texte. Il faut présenter deux copies de cette formule pour chacun des thèmes, la première pour la mise en page de la carte choroplèthe et la seconde pour la mise en page de la carte de répartition par points. Le texte doit être rédigé dans les deux langues officielles et il ne doit pas dépasser les limites indiquées sur la formule. La ponctuation, l'espacement et l'utilisation des lettres majuscules et minuscules doivent être soigneusement établis. Si le texte pour un thème donné ne peut être appliqué à l'ensemble des 12 RMR, on doit remplir une paire de formules distinctes pour chacune des régions.

La formule 3 sert à présenter les données comparatives sur la RMR, la province et le Canada. Dans le cas de l'histogramme cumulé (s'il y a lieu), les données comparatives sont présentées, sous forme de chiffres arrondis relatifs à un maximum de cinq grandes variables, plus les totaux, aux lignes 1 à 5 et à la ligne T. Pour ce qui est du tableau des données comparatives, les pourcentages calculés à partir de données arrondies (ou autres valeurs appropriées, par ex., les revenus moyens) doivent être inscrits à la ligne P.

Après le traitement de ces données d'entrée, l'équipe de production remet les tracés préliminaires, où figurent seulement le texte et les données comparatives, à l'agent spécialisé pour qu'il les examine.

Specification of Text on Maps (See Guide Item 2)

All CMAs :

NOTE ABOUT BONUS LETTERS:

all letters in the alphabet being used are e

on average letter. In preparing text for the

... and you may use one bonus letter space for every

Guide item 2.11)

[illegible]

Continue

Continued on Page 2 ...

☒ All CMA's:
or Specify CMA:

[illegible]

(1)	English French	65 and over 65 et plus
(2)	English French	20-64 20-64
(3)	English French	5-19 5-19
(4)	English French	under 5 moins de 5
(5)	English French	
T)	English French	population population

For use by Subject matter officer(s):.

- 1 Form completed by: _____ date: _____
- 2 Form approved and forwarded to GCG by: _____ date: _____
- 3 _____ date: _____
- 4 _____ date: _____
- 5 Plots received and text verified by: _____ date: _____
- 6 Plots rejected - see memo attached: _____ date: _____
- 7 Plots accepted - GCG notified by: _____ date: _____

For use by GCG production team:

- 3 Form received by: _____ date: _____
- 4 Text coded and job submitted by: _____ date: _____
- 5 Plots checked and forwarded to subject matter by: _____ date: _____
- 8 Plots rejected - corrective action taken by: _____ date: _____
- 10 Plots accepted - files stored and documented by: _____ date: _____

Specification of Comparative Data (see Guide Item 2)

☒

or specify CMA

Specify: HIGHLIGHT Row # 4.

Province																																				
Saskatchewan	1	1	6	1	7	0	5	1	8	5	9	0	2	5	2	1	3	0	8	1	4	2	0	9	6	8	3	1	3							
Alberta	8	8	4	9	1	6	3	3	8	5	1	3	1	8	2	3	5	5	6	8	1	3	5	1	8	7	9	6	5	2	2	3	7	7	2	4
Alberta	8	8	4	9	same data as above																															
British Columbia	2	9	8	1	7	5	1	6	1	8	7	0	0	6	3	6	3	8	0	1	9	1	2	0	5	2	7	4	4	4	6	7				
	7	0	0	0																																

For use by subject matter officer(s)

I. Form completed by:

2. Form approved and forwarded to GCG by:

6 Plots received and data verified by:

7 Plots rejected - see memo attached

9. Plots accepted - GCG notified by:

For use by GCG production team:

3. Form received by:

4. Data coded and job submitted by:

5. Plots checked and forwarded to subject.
matter by: _____ date: _____

8. Plots rejected - corrective action taken: _____ date: _____

10. Plots accepted - files stored and documented by: _____ date: _____

5.4 Selection of Class Intervals and Dot Rate

The class intervals for the choropleth map and histogram(s) and the dot rate for the dot distribution map are supplied by the subject-matter officer on Form 4 (Figure 12). Refer to Chapter 3 and Items 2.5, 2.7, 2.8, and 2.11 of this guide for an explanation of the requirements at this stage. Use a separate form for each CMA.

After processing this input, the production team will provide copies of the final plotted maps to the subject-matter officer for verification. This is the last verification step before printing, so review the plots carefully.

5.4 Sélection des intervalles de classe et de la densité des points

L'agent spécialisé inscrit sur la formule 4 (voir figure 12) les renseignements concernant les intervalles de classe pour les cartes choroplèthes et les histogrammes, et la densité des points pour la carte de répartition par points. Pour une explication des exigences particulières de cette étape, se reporter au chapitre 3 ainsi qu'aux points 2.5, 2.7, 2.8 et 2.11 du présent guide. Une formule distincte doit être établie pour chaque RMR.

Après le traitement de ces données d'entrée, l'équipe de production remet des exemplaires du tracé définitif des cartes à l'agent spécialisé pour qu'il les examine. Comme il s'agit de la dernière étape de vérification avant l'impression finale, les tracés doivent être soigneusement passés en revue.

Metropolitan Atlas Series

FORM 4

Specification of Class Intervals and Dot Rate (see Guide item 3)

Theme No: 3 Theme Name: Population under five years

Specify:

☒ Layout 1
☐ Layout 2☒ Dot Map
☒ Choropleth MapSpecify CMA: Edmonton

1. Specify Class Intervals for choropleth map, Indicate %, \$, + or - sign, etc. as appropriate.
- ☐ 3 classes (1) less than / moins de 2.5%
☐ 4 classes (2) 2.5% to / à 7.5%
☒ 5 classes (3) 7.5% to / à 12.5%
(4) 12.5% to / à 17.5%
(5) 17.5% or more / ou plus
- Note: The intermediate class(es) should be written as "X% to Y%". The top and bottom classes may take the same form, or alternatively be written as "Y% or more" and "less than X%".
2. Specify Frequency Histogram Intervals for choropleth map page From 0% to 20% by intervals of 2.5% item
3. Specify Dot Rate for dot map: One dot represents persons (persons, dwellings, households, families)
 Un point représente personnes (personnes, logements, ménages, familles)
4. Specify Frequency Histogram Intervals for dot map page - LAYOUT 2 ONLY: From to by intervals of .

For use by GCG production team:

1. SAS (UNIVARIATE, FASTCLUS, and DISCRIM procedures) jobs submitted by: Date:
2. GDMS and CHANG jobs submitted by: Date:
3. SAS, GDMS and CHANG printouts checked and forwarded to subject matter by: Date:
6. Form received, GIMMS commands coded and job submitted by: Date:
7. GIMMS printouts received and verified, Plot request submitted by: Date:
8. Plots received, checked, and copies sent to subject matter by: Date:
11. Plots rejected - corrective action taken by: Date:
13. Plots approved - originals filed in climate - controlled storage by: Date:

For use by subject matter officers:

4. Printouts received, and data analysed. Dot rate, Class Intervals and Frequency Histogram Intervals indicated on this form by: Date:
5. Form approved and forwarded to GCG by: Date:
9. Plots received and verified by: Date:
10. Plots rejected - see comments on plot copies: Date:
12. Plots approved - GCG notified by: Date:

